1954

A study of pulse delta modulation.

Armstrong, William Floyd

Monterey, California: U.S. Naval Postgraduate School

A STUDY OF PULSE DELTA MODULATION

WILLIAM FLOYD ARMSTRONG Jr.

A STUDY OF PULSE DELTA MODULATION

W. F. Armstrong, Jr.

A STUDY OF PULSE DELTA MODULATION

by

William Floyd Armstrong, Jr.
Lieutenant, United States Navy

Submitted in partial fulfillment
of the requirements
for the degree of
MASTER OF SCIENCE
IN
ENGINEERING ELECTRONICS

United States Naval Postgraduate School
Monterey, California
1954

This work is accepted as fulfilling

the thesis requirements for the degree of

MASTER OF SCIENCE

IN

ENGINEERING ELECTRONICS

from the

United States Naval Postgraduate School

PREFACE

This work on Delta Modulation was accomplished in major part at the U. S. Naval Postgraduate School, and was executed for the purpose of compiling useful information about the technique into a compact form. Certain portions of this work have not previously been set down to the author's knowledge, and it is hoped they may contribute in some way to a furtherance of the art.

Acknowledgement is gratefully made to Dr. [illegible]. M. Bauer, Professor of Engineering Electronics at the U. S. Naval Postgraduate School for his advice and assistance. Appreciation is also expressed to Mr. S. Lewinter, of Federal Telecommunications Laboratories, for his cooperation and advice in the construction of the dual polarity pulse system.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

CHAPTER I

Introduction

For more than a decade the use of pulse modulation techniques for communications transmission have been in use in many and varied forms. These forms were based on the principle that any sine wave could be reproduced if the amplitude of the wave was known at each of two equally spaced time intervals of each cycle. Thus, by transmitting this amplitude information at a rate equal to at least two times the highest frequency involved, the transmission of any one communications channel could be interlaced with many others, since only a very small fraction of a complete cycle was needed to transmit all the necessary information about that channel to allow it to be reconstructed in its entirety at some receiving point. This interlacing of many communications channels to be transmitted over a single wire or at a single carrier frequency makes up what is known as the pulse modulation multiplex system.

This paper is not to deal with the multiplex problem itself, but rather with the methods of converting an audio input wave into a group of pulses which contain sufficient information to allow the complete audio wave to be reconstructed, and in particular with the method known as delta modulation to be discussed more fully in succeeding chapters.

The amplitude information to be transmitted can be contained in a pulse in several forms. If the pulse amplitude itself is varied as a function of the input signal amplitude, we have what is known as pulse amplitude modulation (PAM). In contrast to this, the method known as pulse width modulation (PWM) has the signal amplitude information contained in

the width of the transmitted pulse, the widths being greater or smaller than some fixed mean width to indicate a signal amplitude which is larger or smaller than some fixed mean amplitude. A third method, known as pulse position modulation (PPM), transmits the required amplitude information by shifting the time position of each pulse to a slightly earlier or later time, thus varying it about the mean instant of sampling for that channel. The time of occurrence about this mean time would be a function of the input signal amplitude at the instant of sampling.

The three systems outlined thus far have one common shortcoming. They are susceptible to noise interference in transmission when signal to noise ratio is low which can produce a considerable amount of distortion in the reconstructed signal outputs. The pulse amplitude method could have a pulse amplitude, particularly one representing a small amplitude, altered by a noise burst, which would result in erroneous amplitude information at the receiver. So also could the pulse width be apparently widened by a burst of noise occurring just prior to or at the end of a transmitted pulse; likewise, the pulse position could be apparently shifted to an earlier instant by a noise pulse occurring just prior to the desired pulse itself.

To overcome this shortcoming and permit a more satisfactory signal reconstruction for low signal to noise ratios, the pulse code systems were devised which were less susceptible to interference in that the signal amplitude information required was contained in the mere presence or absence of the pulses and not in any characteristic of the pulses themselves as to width, amplitude, position, etc., Delta modulation is of this latter type.

CHAPTER II

Pulse Code Modulation

Pulse code modulation was devised as a means of utilizing pulse techniques in the presence of noise of such magnitude as to cause untolerable distortion in other pulse systems. It can be shown that a signal can be transmitted with a signal to noise ratio as low as 2:1 and be reconstructed at the receiver as if the transmission had occurred at a ratio equal to infinity, if the noise generated in the receiver itself is ignored, as well as the quantizing noise which is inherent in a code system.

The fore-runner of pulse code modulation was a method known as pulse numbers modulation (PNM), where at each sampling instant the signal amplitude was transmitted as a series of pulses, rather than a lone pulse, and the number of pulses transmitted was a measure of the amplitude. Thus the receiver had only to determine how many pulses were present at each sampling instant. It is obvious that the number of pulses that could be included in any one sampling is limited both by the sampling rate and also by the number of channels multiplexed into one system on a time division basis. The duty time of any one channel could be defined as the fractional part of a second the channel has available to it for making a sampled transmission, or

$$D = \frac{1}{(N+1)} \text{ sec.}$$

where N is the number of information channels multiplexed, and the added one in the denominator is a separate synchronizing channel to synchronize the division of received pulses at the receiver. In some systems, this

synchronizing signal is not allotted a separate channel, but instead is added into one of the regular channels on an opportunity basis. Within the duty time of the channel, it will be necessary to make at least $2f_h$ transmissions of pulse groups, leaving a transmission time per pulse group of

$$T_g = \frac{1}{2 f_h (N+1)} \text{ sec.}$$

where, T_g - Transmission time per pulse group in seconds.

f_h - Highest frequency input signal to any channel.

In most communications transmissions in the audio range, the audio frequency band to be transmitted for a good degree of intelligence is 200-3500 cps. In an eight channel multiplex system where f_h is 3500 cps,

$$T_g = \frac{1}{2 \times 3500 (8+1)} = 16.15 \ \mu \text{ sec.}$$

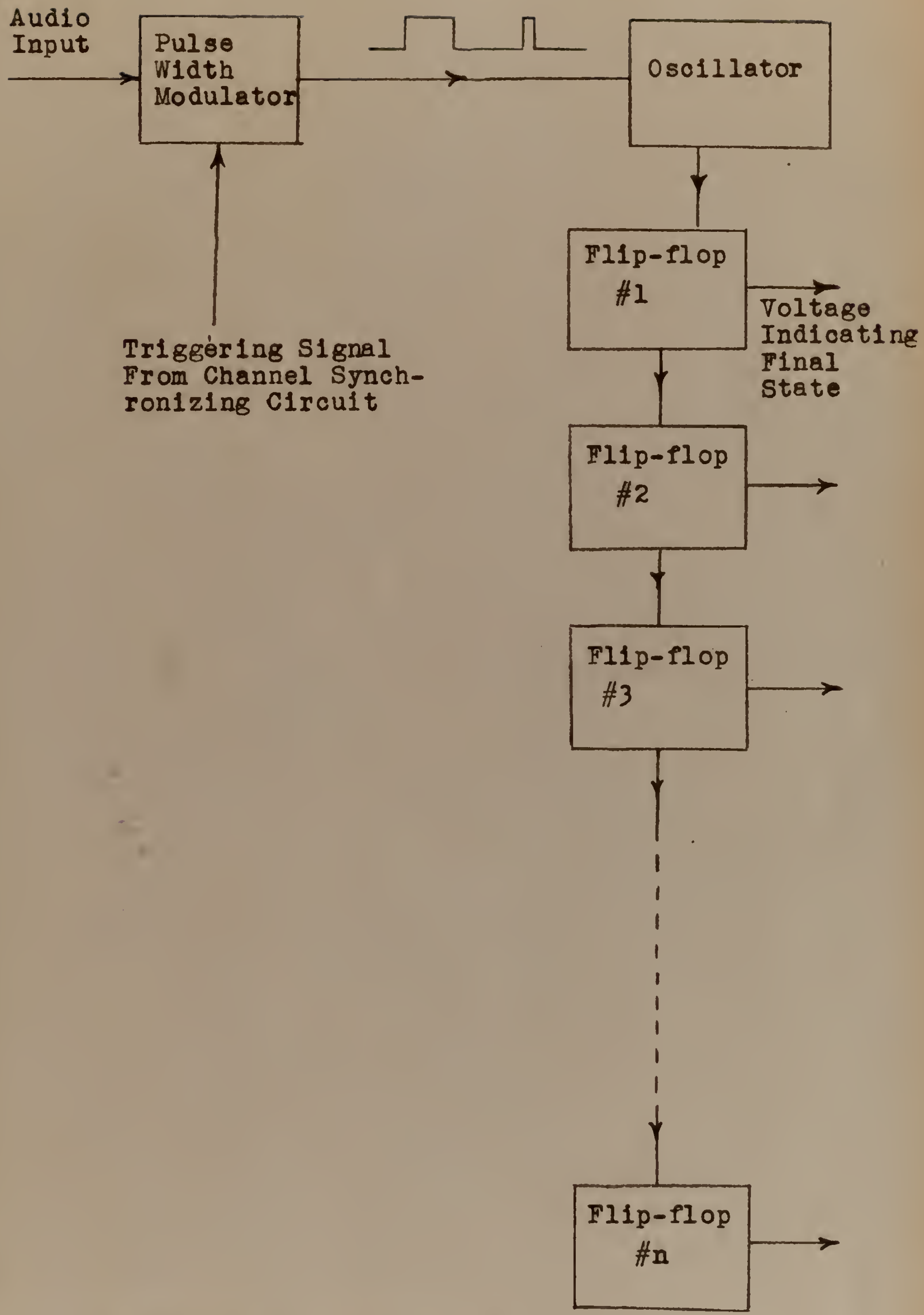
From this we see that if each pulse is of 1 u second duration and if there is 1 u second between pulses, this type of system would provide for only eight discreet levels in the amplitude quantizing of the audio signal. This would result in a large amount of distortion in the reconstructed signal. This distortion would be due to the finite number of measuring levels available to represent the signal amplitude and is commonly referred to as quantizing noise, since it results from the quantizing measurement and representation of the signal amplitude.

The true pulse code modulation system improved on this numbers system by using a binary code and thus allowing the absence of a pulse to carry information as well as the presence of a pulse. If a code group of n pulses is available to represent the signal amplitude for one sampling, then the

number of discreet levels available for quantizing the signal amplitude would be 2^n, since there would be that many pulse group combinations available on a binary basis. Thus for n=5, 32 levels would be available, whereas for n=7, 128 levels could be used. The five and seven digit codes are of most common usage.

At each instant of sampling in a pulse code system, the system must measure the input signal amplitude, quantize it into one of 2^n levels, and transmit one group of up to n pulses representing this level. The quantizing of the instantaneous input signal amplitude is ordinarily done by one of three coding methods; by a counting circuit, by use of a special coding tube devised by Bell System Laboratories, or by use of a feedback comparison method.

Encoding by use of a counting circuit can be most readily understood by reference to Fig. 1. Here the signal is fed into a modulator which generates a positive gate whose width is varied as a function of signal amplitude, the larger the positive amplitude the wider the gate with the largest negative amplitude producing the narrowest gate. This positive gate permits the output of an oscillator to be applied to the first of a series of n bi-stable flip-flop circuits, so tied together that each time one flip-flop is triggered twice, it triggers the following flip-flop once. The oscillator must have such a frequency that it can produce 2^n cycles during the period of the maximum gate applied to it, which in turn represents the maximum input signal amplitude capable of being quantized. As soon as the positive gate disappears, the first flip-flop will no longer receive triggering impulses from the oscillator and all n flip-flops will remain as last triggered. The flip-flops will be in one of two states.

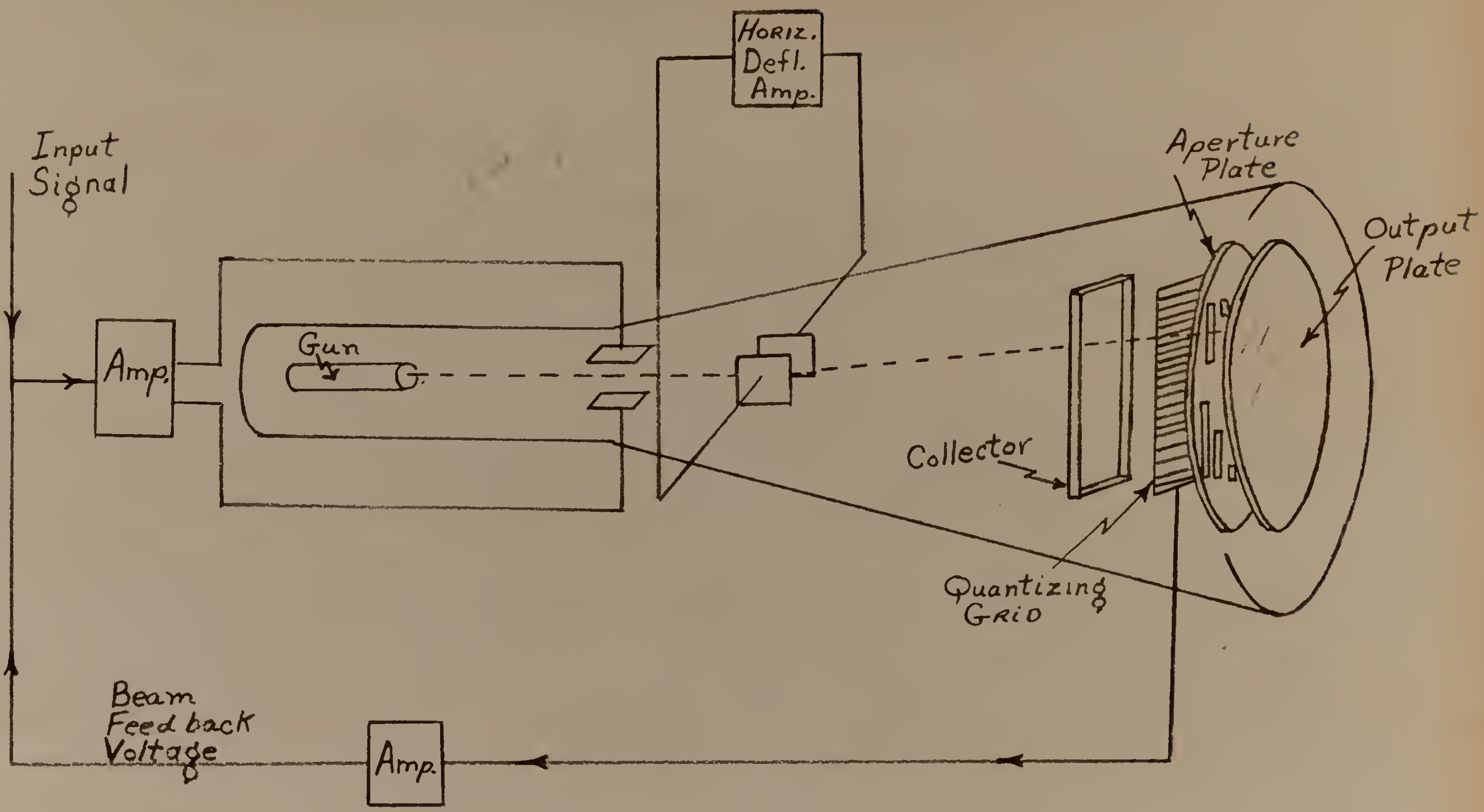


Counting Method for PCM Encoding

Figure (1)

By taking an output from the plate of the same tube in each flip-flop, the voltage out will be an indication of which state each particular flip-flop ended in. If the circuit is in state one, a pulse will be placed in the n digit code group in the same relative position as the flip-flop is in the entire chain. If the circuit is in state two, a pulse will be omitted from this position in the group. This counting occurs just prior to the instant of code group transmission, as the n flip-flops must be allowed to stabilize, and then each one in turn permitted to add or omit a pulse to the code group.

Bell Laboratories devised a coding tube, a modified cathode ray tube, which performs the coding operation in one horizontal sweep. In this tube, shown in Fig. 2, a perforated mask lies between the deflection plates and the anode. The beam of electrons is displaced vertically an amount proportionate to the input signal amplitude. At the instant of sampling, a horizontal sweep voltage is applied, and as the beam moves across the tube, it strikes a certain combination of apertures in the mask, depending upon its vertical position. In all, there will be 2^n different aperture combinations to be found during the horizontal sweeps. As the beam strikes an aperture, electrons will flow through the mask to the anode, and a pulse of current will flow in the output circuit. To prevent the vertical position of the beam from changing during any one sweep, a quantizing grid of 2^n horizontal wires is placed immediately in front of the coding mask. If the beam attempts to alter its vertical position during a horizontal sweep, beam electrons will strike one of the grid wires which emit secondary electrons. These are in turn caught on a positively biased collector, causing current to flow in the quantizing feedback circuit.

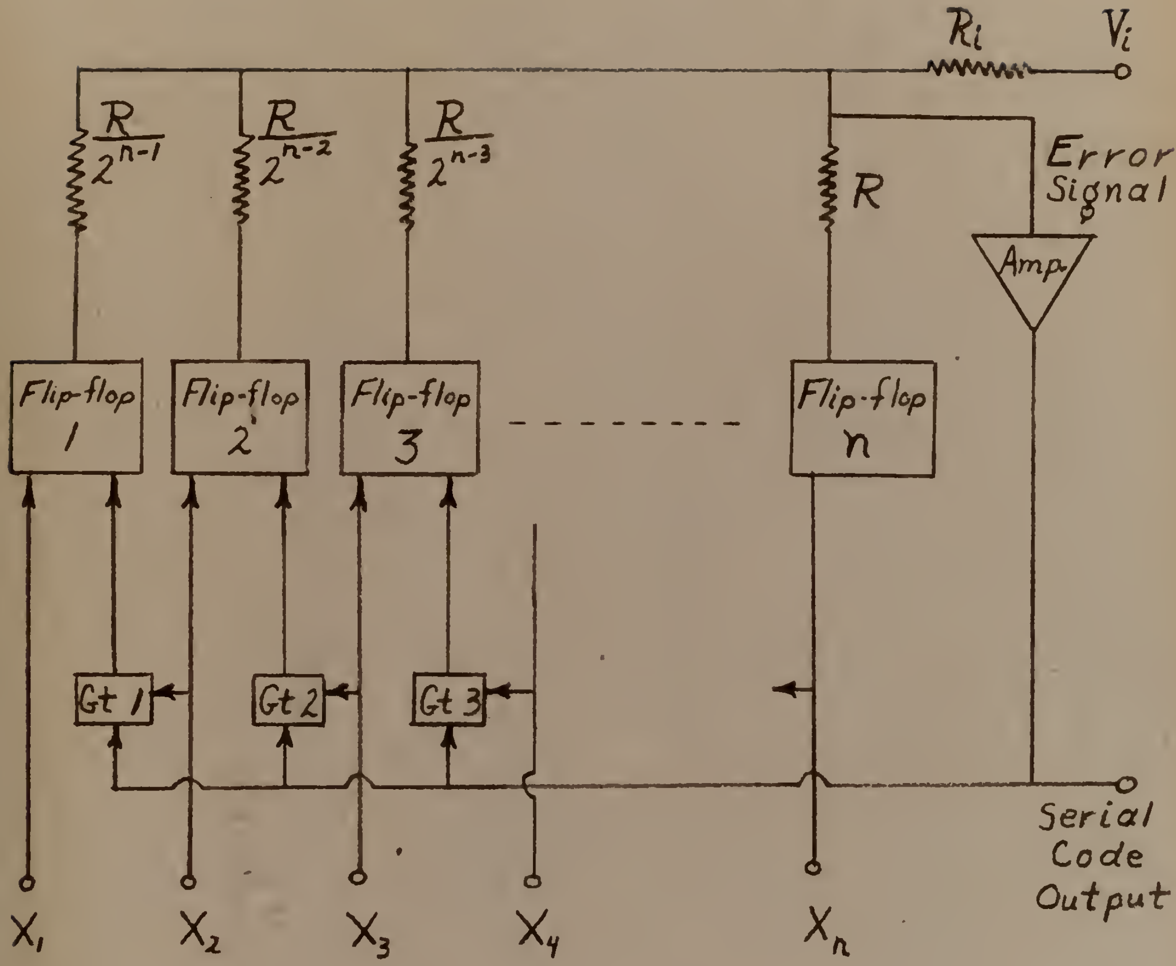


Coding Tube For PCM

Figure 2

This excitation in the quantizing circuit is fed back to the vertical deflection plates with such polarity as to cause the beam to be deflected back away from the quantizing grids upon which it is impinging and thus reduces the current flow in the quantizing circuit to a minimum. Since the flow of current in the quantizing feedback circuit will always be unidirectional, a quantizing bias is used which tends to displace the beam upward, toward a higher quantizing wire. This allows the feedback circuit to always produce a signal to the vertical plates which is of such a polarity as to lower the beam. This feature insures that the correction of vertical position will always be in the correct direction. It can be looked upon as if the quantizing bias pushes the beam up toward the next highest wire, while the feedback correction tends to push it down to some equilibrium position. This process results in the beam making a complete horizontal sweep at a constant vertical level set by the signal amplitude at the instant the sweep starts.

A third method of coding utilizes a feedback comparison process. Smith (11) has very adequately described the method by use of Fig. 3. The n digits of the code group, giving a binary representation of the sampled amplitude, are each determined by the n bi-stable flip-flops, each of which has a control gate shown. By this feedback method, the most significant digit of the n digits is determined first. A source of equally spaced timing pulses is required for inputs X_1, X_2, etc., with X_1 occurring first. Initially, pulse X_1 will set all flip-flops to state 1, except flip-flop #1 which will be set to state two. The resistors in series with the output of each flip-flop make up a decoding network. The resistors are in the ratio 1, $\frac{1}{2}$, $\frac{1}{4}$, 1/8, etc., with the smallest resistor associated with



Feedback Method For PCM Encoding

Figure 3

the flip-flop representing the most significant digit, or flip-flop #1 in this case. When any flip-flop is in state one, no voltage is applied at its output to the decoding resistor in series with it. Since each flip-flop will produce the same output voltage in state two, the voltage each is able to deliver to the input of the error amplifier depends on the voltage drop in its associated decoding resistor. Thus the larger the decoding resistor, the smaller is the voltage fed back to the input of the error amplifier.

When the timing pulse is applied at X_1, only flip-flop #1 will be in state two, which in turn feeds a voltage into the input of the error amplifier; the error amplifier is also receiving the signal voltage. This error amplifier responds to the difference of these two inputs. If the error signal is negative (signal voltage larger than voltage fed back from flip-flop #1) no action results and flip-flop #1 is left in state two. If the error signal is positive, the amplifier produces a positive gate in the serial code output, which also gates on Gt 1. With Gt 1 gated on, timing pulse X_2 resets flip-flop #1 to state one. Thus reset, it does not apply any feedback to the decoding circuit. When timing pulse X_2 is applied, it sets flip-flop #2 to state two, permitting it to feed back a voltage to the error amplifier input, and the action is as before. Each time the error voltage is positive, the last triggered flip-flop is reset to state one where it produces no feedback. Since each successive flip-flop produces less and less feedback, the result is that at the input to the error amplifier the error voltage will increase by steps until it becomes positive. It will then be reduced by an amount equal to the last added step and immediately increased by a smaller step. Each time the

feedback is reduced, a pulse is placed in the serial code output. Thus a signal of zero amplitude would have a serial code output of n pulses, since zero feedback would be required to prevent a positive error voltage.

The three coding systems described, are each representative of a specific coding method. The last mentioned system is the forerunner of pulse delta modulation, which is basicly a feedback comparison system using a one digit code. This delta system is to be the subject of the remainder of this report.

CHAPTER III

Description of Delta Modulation

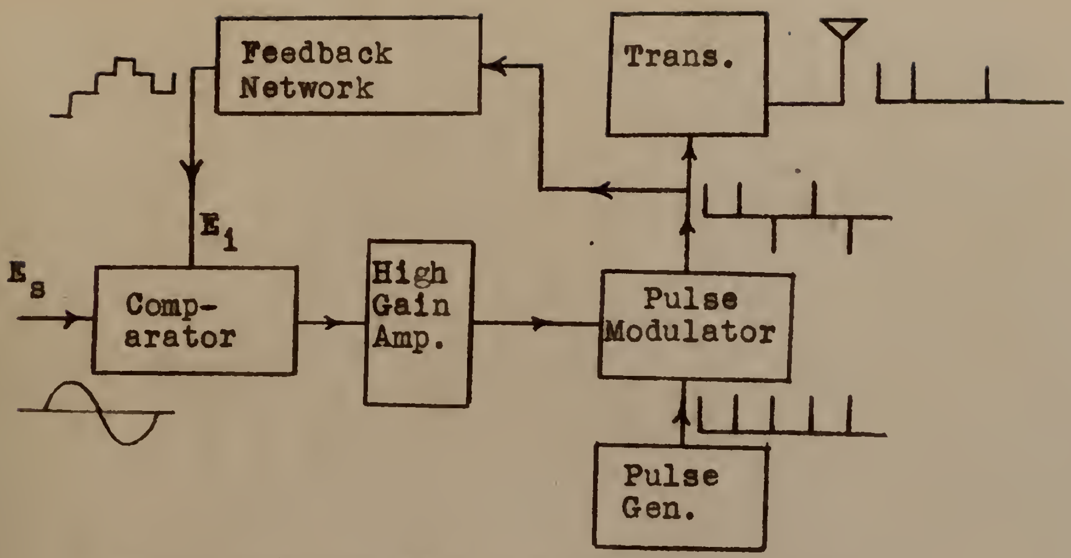
Delta modulation is the newest innovation in the field of pulse modulation. Basicly, the system is of the pulse code type, using a feedback comparison technique for control of the circuit operation. The code itself could only be called a uni-digit binary code, since each code group consists of the presence or absence of a lone pulse. Unlike the other forms of pulse code modulation mentioned previously, the information contained in the lone pulse does not contain the quantized amplitude of the input signal. Rather, the pulse merely indicates that the instantaneous input signal amplitude at the time of sampling is greater or less than an approximating image signal constructed by the pulses transmitted. This image signal is constructed at the transmitter for comparison purposes, and at the receiver it constitutes the received information after being passed through a low pass filter. Since each pulse contains so little information, a much higher sampling rate is needed than in conventional pulse code modulation where the rate can be as low as $2f_h$.

Figure 4(a) is a block diagram of a delta modulation system. The input signal E_s is placed in a comparator circuit where it is compared with a signal E_i which is a reconstructed image of the input E_s. The output of this comparator will be difference voltage of positive or negative polarity, depending upon which of the two voltages compared is larger. This difference voltage is amplified and used to modulate the pulse modulator, which receives pulses at a fixed rate from the pulse generator. If the input signal voltage is larger than the reconstructed image voltage

E_i, an output of positive polarity will come from the comparator and following amplifiers to cause the modulator to emit a positive pulse which will be used in the feedback network to increase the image signal voltage. If the input signal voltage is less than the image voltage, a negative polarized signal will reach the modulator, causing it to emit a negative pulse which in turn will be used to decrease the image voltage feedback to the comparator. The chain of positive and negative pulses is fed to a transmitter which will eliminate the negative pulses and transmit the positive ones. At the receiver, the absence of a pulse is interpreted as a negative pulse. This same chain of pulses passes through the feedback net as previously mentioned. This network reconstructs the input signal from the positive and negative pulses and sends the image to the comparator for comparison with the original.

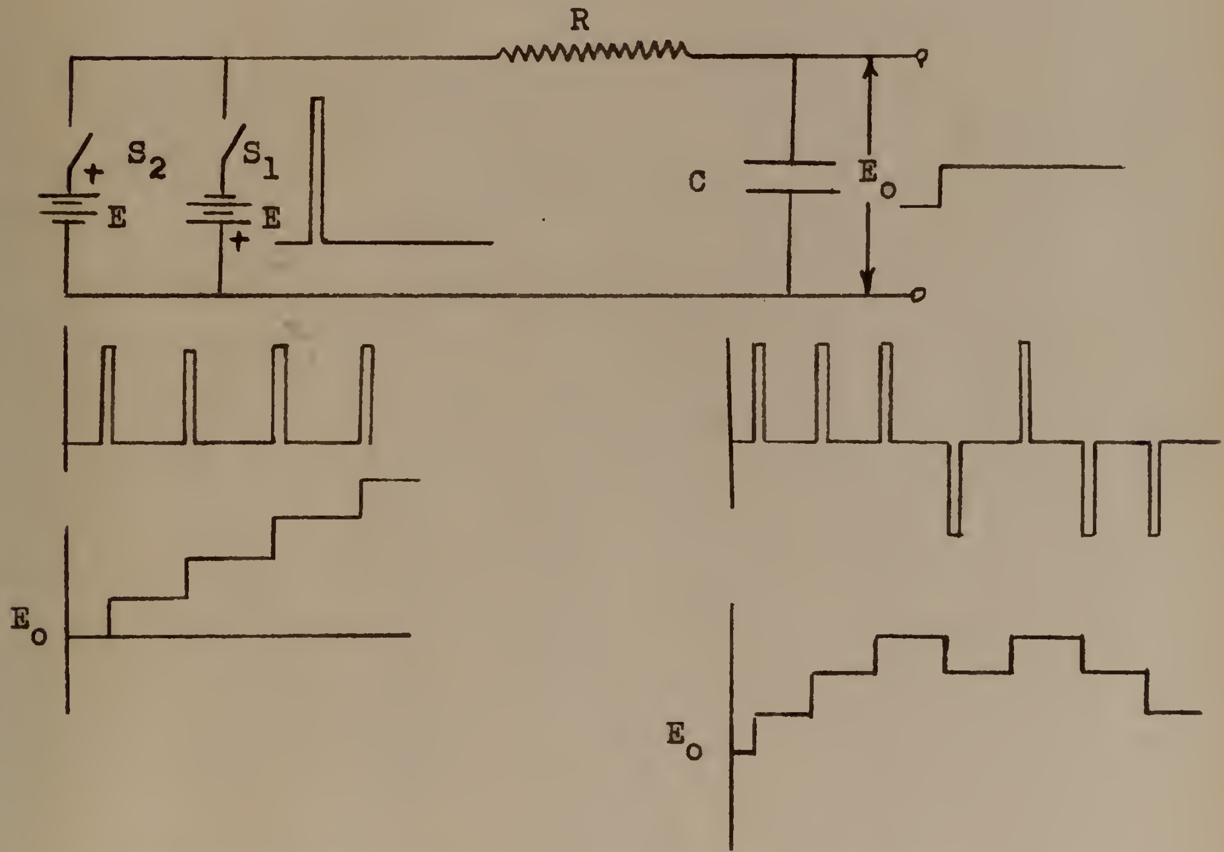
The heart of the decoder in the feedback loop is an RC network, the operation of which is as shown in Fig. 4(b). If S_2 is closed for a short period of time, say $1\mu s$, a positive voltage pulse will be applied to the RC series circuit. If the time constant of this circuit is much much larger than $1\mu s$, the output response to the pulse will be approximately a step function with an amplitude that is small compared to the pulse amplitude, as shown. If S_2 is closed for very short periods at repeated and equally spaced intervals, the voltage across C would be as shown in the lower portion of the figure. Likewise, if S_1 is closed so as to deliver short negative pulses, the voltage across C will drop one step each pulse, and a series of positive and negative pulses might produce a wave form such as shown in Fig. 4(b).

If now the input signal is a sine wave, the reconstructed image



Block Diagram of Delta Modulation System

(a)



Decoding Network and Output For Pulses Shown

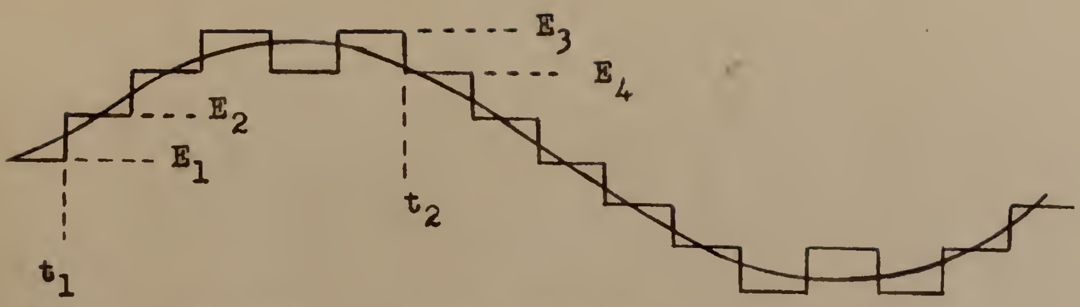
(b)

Figure 4

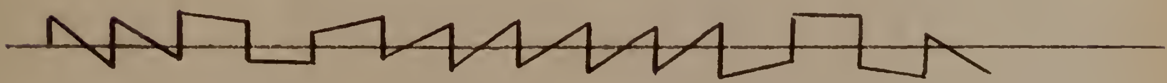
would be a stepped voltage wave, shown superimposed on the sine wave in Fig. 5(a). The comparator would act only on the difference signal shown in Fig. 5(b) and the pulse modulator would emit pulses as shown in Fig. 5(c) to build up the stepped voltage wave. At time t_1, since the stepped wave is less than the sine wave, a positive pulse is emitted and the stepped wave is increased from E_1 to E_2. Likewise at time t_2 the stepped voltage is larger than the input sine wave, and a negative pulse is emitted and the stepped voltage wave is reduced from level E_3 to E_4. By comparing the superimposed wave forms of Fig. 5(a), one sees that each time a pulse is added, it is added in such polarity as to move the stepped voltage wave toward the sine wave.

If this stepped sine wave from the RC decoding net were also passed through a low pass filter, the result would be a very close approximation to the original sine wave input. This is essentially how the receiver operates. It utilizes an RC decoding network, passes the stepped output through a low pass filter with a cut off frequency of about 3500 cps and amplifies the output as necessary.

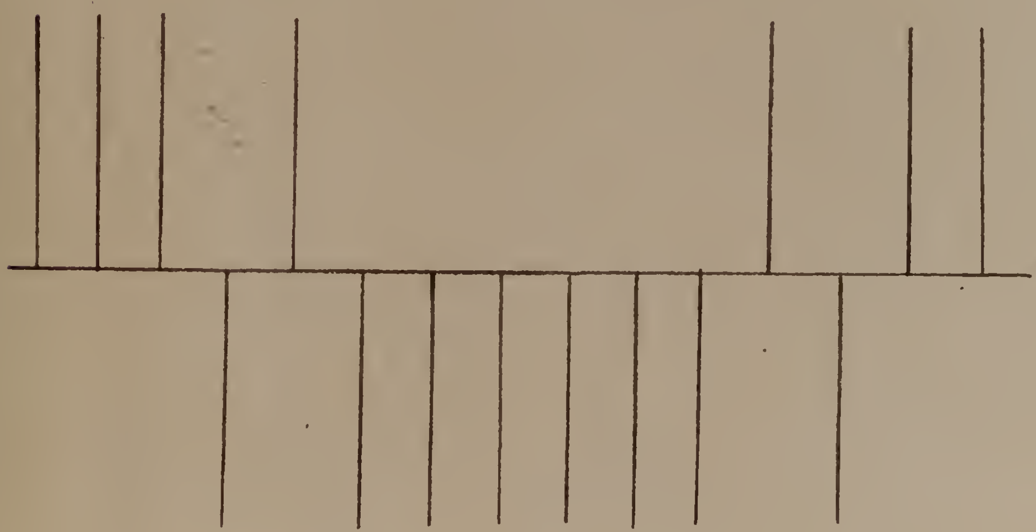
The operation just described constitutes a method utilizing a single RC integrating network in the feedback decoding network. It could also have functioned with a double integrating network in this circuit.



Superimposed Signal and Stepped Image
(a)



Error Voltage
(b)



Resultant Pulse Train
(c)

Figure (5)

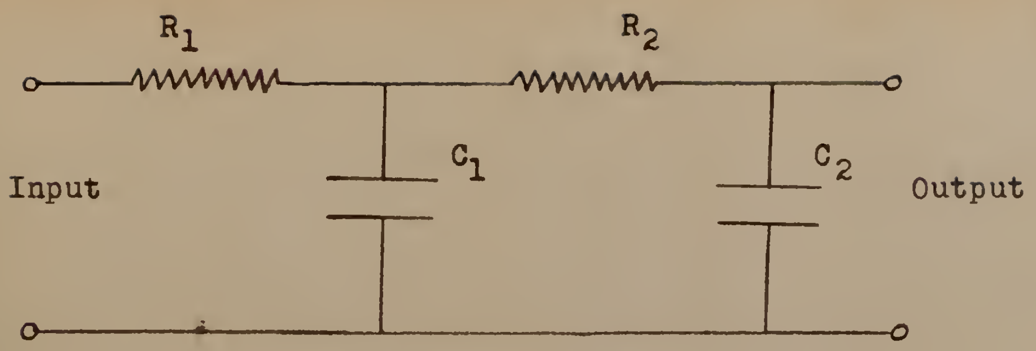
CHAPTER IV

Double Integration Network for Decoding

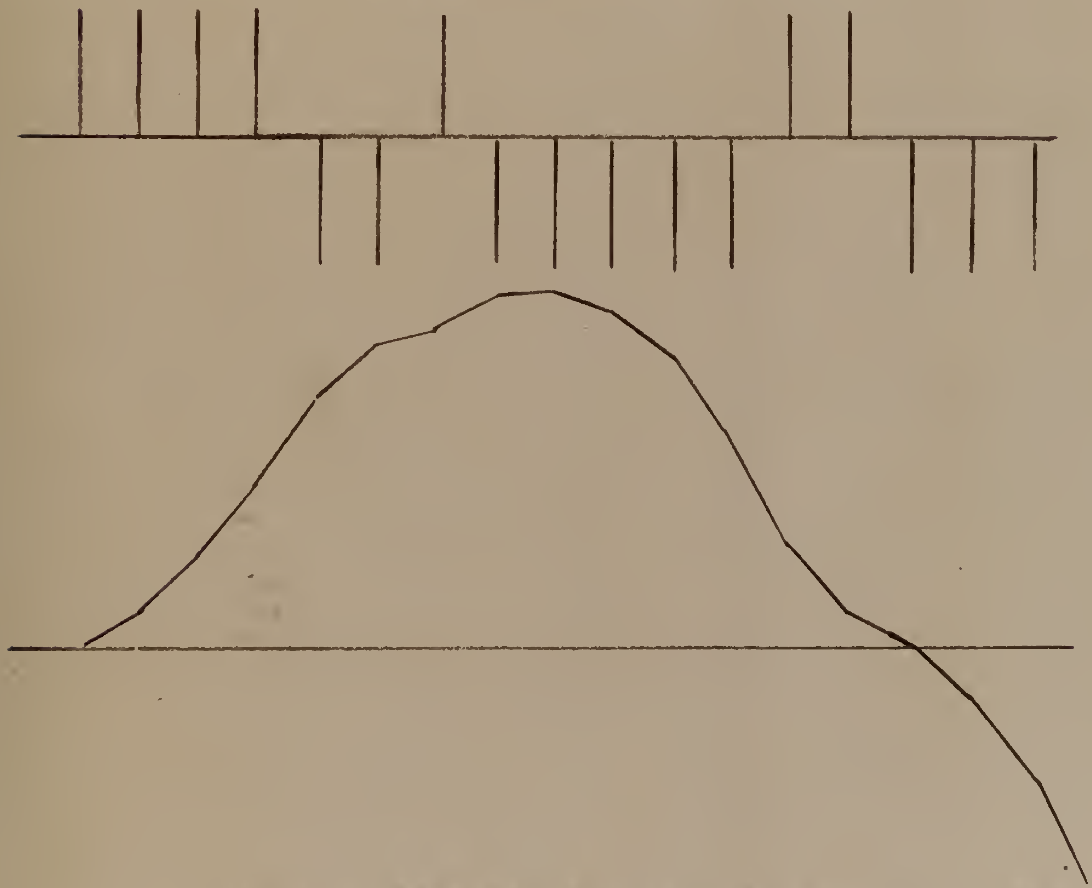
Instead of the RC network previously described as usable in the decoding feedback circuit, one such as depicted in Fig. (6) could also have been used.

Here, the series of positive and negative pulses are applied to the input terminals. If the pulse duration is τ, and R_1C_1 is much much greater than τ, and R_2C_2 is much larger than R_1C_1, then during the application of the pulses, C_1 will charge up in almost a vertical voltage rise, since as before it is using only a very small portion of one time constant. Between pulses, capacitor C_1 will charge up C_2 through R_2. The output voltage across C_2 will rise with a slope fixed by the magnitude of the voltage on C_1. The greater the voltage on C_1, the steeper will be the slope of the voltage output of C_2. Each time a positive pulse is applied to the network, the voltage across C_1 is increased one step, just as for a single integrating network, and since C_2 will then be charging to a higher value than before, the slope of the output voltage will increase. Conversely, each time a negative pulse is applied to the network the voltage across C_1 is reduced, and the slope of the output voltage from C_2 decreases, resulting in an effect shown in Fig. (6b). We see here that the pulses control the slope of the image voltage rather than the stepped level as in the case of the single integrating network.

In the use of this type circuit, a problem of stability arises in that for a zero level input signal, the output may cross the zero position with a rather steep slope. This would cause the network to put out a



Double Integration Network
(a)



Input Pulse Train and Vari-sloped Output Wave
(b)

Figure (6)

lengthy group of pulses of the proper polarity to bring the output back to the zero level. These many pulses would cause the output voltage to recross the zero level at an equally great or greater slope than before and an instability could easily arise which would lead to oscillation of the output.

DeJager has outlined a means of eliminating this oscillatory tendency and increasing the effectiveness of the circuit by use of a circuit such as shown in Fig. (7). Here the voltage across C_2 has added to it the voltage across R_3 to make up the output of the network. This is a type of predictor circuit, since this output is indicative of what the voltage across C_2 will be at some future time $(t+\Delta t)$, assuming a constant slope is maintained until the next pulse is applied.

Thus,

$$e_o(t) = e_2(t+\Delta t)$$

If $\frac{de_2}{dt}$ is the slope of the output voltage between the time of two pulses,

$$e_o(t) = e_2 + \Delta t \frac{de_2}{dt} \qquad (1)$$

Letting i be the current through R_3 and C_2,

$$e_2 = \frac{\int i\,dt}{C_2}$$

$$\frac{de_2}{dt} = \frac{i}{C_2}$$

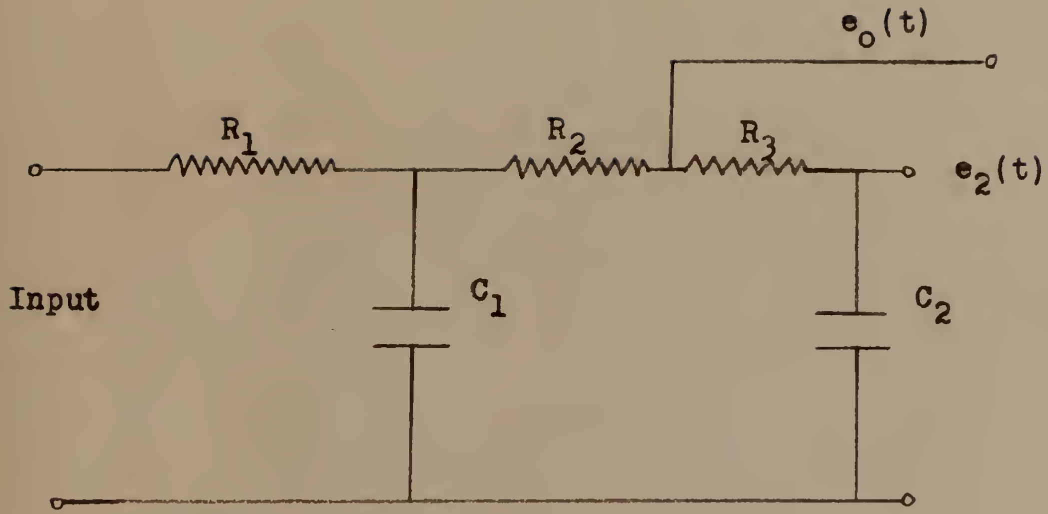
and substituting into (1), we get

$$e_o(t) = e_2 + \Delta t \frac{i}{C_2} \qquad (2)$$

The difference of potential across R_3 can be expressed as,

$$i R_3 = e_o(t) - e_2(t)$$

Also, for (2), $e_o(t) - e_2(t) = \Delta t \frac{i}{C_2}$



Double Integrating Network Using Prediction Voltage Component in Output.

Figure (7)

Therefore, $\Delta t \frac{i}{C_2} = i R_3$

and, $\Delta t = R_3 C_2$ (3)

From equation (3), we see we can regulate the effectiveness of the prediction feature by regulating the R_3C_2 product. An optimum value of about $\frac{1}{f_p}$ for this value is suggested by DeJager (3), f_p being the pulse recurrence frequency.

By using the output $e_o(t)$ in the comparator circuit, we can make a correction or apply a pulse to correct a condition that will exist at some future time $(t+\Delta t)$. If at this future time, the image signal will be greater than the input signal, a negative pulse will be applied, whereas if the image signal will be less than the signal a positive pulse is applied to increase the slope, thus initiating a correction before the error has time to become exaggerated. This technique has eliminated the stability problem of the double integrating network.

CHAPTER V

Technique To Utilize Uni-polarity Pulses

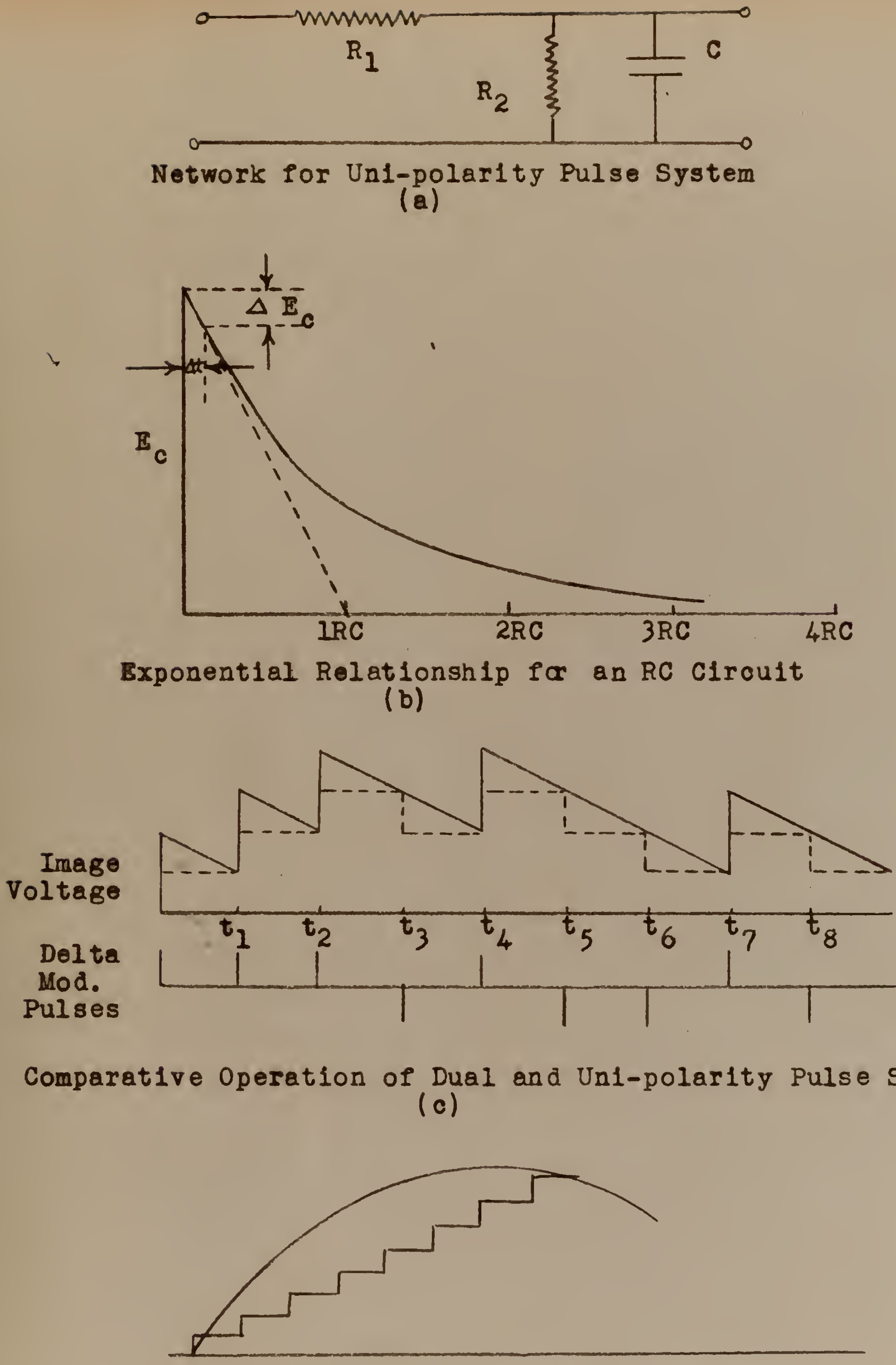
In the discussion thus far, the reconstructed image has been assumed to have been fabricated by the use of positive and negative pulses. This has been done for ease of explanation since the image can be just as easily made up by the use of positive, or negative, pulses alone. This technique adds even more to the simplicity of the equipment involved, and materially reduces the required circuitry at both the sending and receiving points.

The question at once arises as to how the stepped voltage wave could ever have a step down occur if pulses of the opposite polarity are not to be applied. The answer is a shunting device across the integrating capacitor, such as a resistor as shown in Figure 8(a). If the time constant R_2C is much much larger than T_p, the output voltage will decrease after each pulse at a nearly linear rate.

Figure 8(b) shows the exponential change occurring in an RC circuit. From this we see that if the time for discharge is very small compared to 1 RC, that the value of ΔE_c can be expressed by

$$\Delta E_c = \frac{\Delta t}{RC} \times E_{max}$$

If R_2C is of such a value that during the period T_p the voltage on C can be reduced by an amount equal to one half the increase applied by a pulse, the voltage on C at a time T_p seconds after a pulse is the same as if the last pulse had increased the voltage by only one half as much and the charge had not leaked off. Furthermore, if no pulse is applied at this instant T_p seconds later, C will continue to discharge and at the end of a period $2T_p$ the charge on C will be at the same value as it would have



Network for Uni-polarity Pulse System
(a)

Exponential Relationship for an RC Circuit
(b)

Comparative Operation of Dual and Uni-polarity Pulse Systems
(c)

Diagonal Distortion due to System Saturation
(d)

Figure (8)

been had a pulse of the opposite polarity been applied at time T_p and the charge then remained fixed. In Figure 8(c) this principle is illustrated graphically. The solid lines represent the actual voltages occurring accross C, while the dotted lines represent simulated positive and negative step voltages. For the solid line, only the positive pulses apply, whereas for the dotted lines, both positive and negative pulses apply. It is obvious that the two waves will have different high frequency components, but since this output is to be passed through a low pass filter and only the fundamental retained, the end result is the same. It is also to be noted that at times t_1, t_2, t_3, etc., the voltages of both waves are the same. This uni-polarity pulse technique greatly simplifies the equipment involved in the modulation process at the transmitter and equally as much in the detection process at the receiver.

Signal Amplitude Limitations

The maximum signal amplitude which can be handled by a delta modulation system is limited by the maximum slope of the signal itself. This limiting point occurs where the maximum slope of the signal is equal to the slope of the image signal determined by two successive pulses of the same polarity. If the signal amplitude exceeds this critical value diagonal distortion sets in. To determine this critical signal amplitude we can assume a sine wave signal of form Asin(wt). For such a signal the slope is,

$$\text{Slope} = \frac{d[A\sin(\omega t)]}{dt} = \omega A \cos(\omega t)$$

which has a maximum value at wt equal to 0 or 2π.

$$\text{Max. slope} = \omega A$$

If σ represents the height of one step, and the sampling rate is f_p, the time between two successive pulses is T_p, and the slope determined by two successive pulses of the same polarity is

$$\frac{\sigma}{T_p} = f_p \sigma$$

At the critical condition, these two slopes are equal, or

$$\omega A_{max} = f_p \sigma$$

$$A_{max} = \frac{f_p \sigma}{\omega} = \frac{f_p \sigma}{2\pi f_s} \qquad (1)$$

where f_s is the frequency of the input signal.

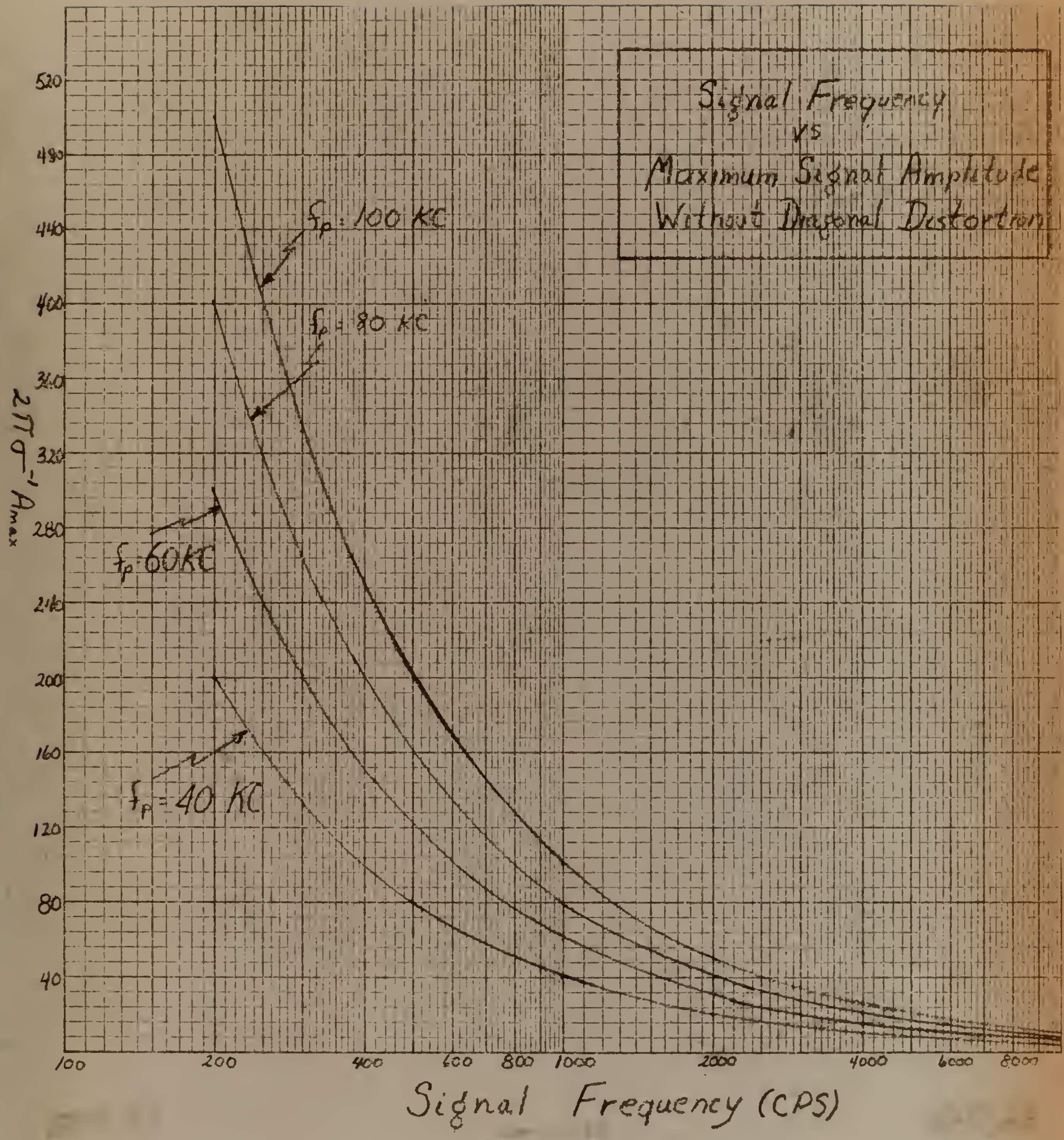
From equation (1) we see that the maximum signal amplitude capable of being encoded without diagonal distortion decreases as signal frequency increases, and increases as either sampling rate or step magnitude increases. Figure (9) is a plot of equation (1) for fixed values of σ . The distortion resulting when A_{max} is greater than $\frac{f_p \sigma}{2\pi f_s}$ is depicted in Figure 8(d).

System Saturation

Saturation of the system sets in when A_{max} exceeds the limiting value given above. To determine the relationship between the signal amplitude which causes complete system saturation and the value given above, let

$$\frac{T_s}{2T_p} = \text{number of pulses available to increase a sine wave at } f_s$$

from minimum to maximum value. If E_s is the peak value of the signal voltage which causes complete saturation, then during the period $\frac{T_s}{2}$ the image signal must be able to have a total increase or decrease of $2E_s$.



Signal Frequency
vs
Maximum Signal Amplitude
Without Diagonal Distortion
f_p : 100 KC
f_p = 80 KC
f_p 60KC
f_p = 40 KC
520
480
440
400
360
320
280
240
200
160
120
80
40
$2\pi\sigma^{-1}A_{max}$
100
200
400
600
800
1000
2000
4000
6000
8000
Signal Frequency (CPS)

Figure 9

Step height = $\frac{\text{total change involved}}{\text{number of pulses producing this change}}$

$$\sigma = \frac{2E_s}{\frac{T_s}{2T_p}}$$

$$\sigma = \frac{4E_s f_s}{f_p}$$

and solving for E_s,

$$E_s = \frac{f_p \sigma}{4 f_s} \qquad (2)$$

Taking the ratio of E_s and A_{max} as given by equations (1) and (2), we get

$$E_s = \frac{2\pi}{4} A_{max} = 1.57\, A_{max}$$

Thus we see that the signal which will completely saturate a system is 1.57 times the signal which just prevents diagonal distortion. This holds for both dual polarity and uni-polarity type systems, where σ for the uni-polarity system will be the residual voltage increase across the storage capacitor T_p seconds after a pulse has been applied.

<u>R_2 and R_1 relationships</u>

To determine the requisite relationships between R_1 and R_2 in the uni-polarity pulse system shown in Figure 8(a), we know that under a condition of complete saturation a decay of $2E_s$ must occur in a time $\frac{T_s}{2}$. Let E_p be the peak pulse amplitude applied and furthermore let E_{dc} be the quiescent DC level of the capacitor C of Figure 8(a). We will also assume that both $(E_p - E_{dc})$ and E_{dc} are much much larger than E_s. Then during the decay period,

$$2E_s = \frac{E_{DC} T_s}{2 R_2 C}$$

$$R_2 C = \frac{E_{DC} T_s}{4E_s} = \frac{E_{DC}}{4E_s f_s} \qquad (3)$$

During a pulse the voltage on the capacitor will increase by 2σ.

$(E_p - E_{dc})$ will be the effective pulse height.

$$2\sigma = \frac{(E_p - E_{DC})\tau}{R_1 C}$$

where τ is the pulse duration. From this,

$$R_1 C = \frac{(E_p - E_{DC})\tau}{2\sigma} \qquad (4)$$

From equation (2), we can derive the expression

$$\sigma = \frac{4 E_s f_s}{f_p} \qquad (5)$$

Replacing the σ in the equation (4), we get

$$R_1 C = \frac{(E_p - E_{DC})\tau f_p}{8 E_s f_s} \qquad (6)$$

Taking the ratio of (3) and (6) we obtain,

$$\frac{R_2 C}{R_1 C} = \frac{\dfrac{E_{DC}}{4 E_s f_s}}{\dfrac{(E_p - E_{DC})\tau f_p}{8 E_s f_s}}$$

$$\frac{R_2}{R_1} = \frac{2 E_{DC}}{(E_p - E_{DC})\tau f_p} \qquad (7)$$

To make the assumption of E_s small compared to $(E_p - E_{dc})$ and E_{dc} have the same effect for both increasing and decreasing voltages on C, we should require that

$$E_p = 2 E_{DC}$$

then,

$$\frac{R_2}{R_1} = \frac{2}{\tau f_p}$$

or,

$$R_2 = \frac{2 R_1}{\tau f_p}$$

This expression shows the relationship of R_1 and R_2 to be dependent only upon the pulse duration and the pulse rate. In this development the shunting effect of R_2 during a pulse has been ignored. Also assumed is that the only discharge path for C is through R_2. Thus R_1 and R_2 have been represented as the total effective values of resistance on charge and discharge of C respectively and may or may not be single, different or even actual physical resistances. That is, R_1 and R_2 each represent what in actuality might be the combination of several resistors, some of which might be common to both R_1 and R_2, or might include such factors as tube plate resistances, etc.

CHAPTER VI

Signal Distortion in Delta Modulation

Signal distortion in delta modulation is due to two specific causes. The first type of distortion is that previously defined as diagonal distortion and is due to the signal having an amplitude greater than $\frac{f_p \sigma}{2\pi f_s}$.

This type of distortion is completely eliminated by reducing signal amplitude to a value less than this limiting figure. If we assume we are operating with signals of small enough amplitude to be free of this type distortion, we will still have what is known as quantizing distortion, which results from insufficient definition of the signal due to the finite size and limited number of steps. Only for a sampling rate of infinity and a step size approaching zero could we be free of this second type of distortion.

The complex wave, or stepped image signal, can be divided up into two parts, the pure signal and the error signal, which is the difference between the pure signal and the stepped image. Referring back to Figure 5(b), we note that the error signal is principally a succession of triangular waves. For the purpose of computing a signal to noise ratio for a single integrating system, we will follow the pattern used by Bennett (2) in his discussion for pulse code modulation and assume the error voltage is a chain of only triangular waves. For a high pulse rate this is a very fair assumption. The period of the triangular waves will be T_p, the pulsing period, and the peak error voltage will vary from $\frac{\sigma}{2}$ to $-\frac{\sigma}{2}$.

The RMS value of such an error wave is:

$$N = \sqrt{\frac{\int_{-\frac{T_p}{2}}^{\frac{T_p}{2}} \left(-\frac{\sigma t}{T_p}\right)^2 dt}{T_p}}$$

$$= \frac{\sigma}{T_p} \sqrt{\frac{\int_{-\frac{T_p}{2}}^{\frac{T_p}{2}} t^2\, dt}{T_p}}$$

$$= \frac{\sigma}{T_p} \sqrt{\frac{T_p^2}{12}} = \sigma/\sqrt{12} \qquad (1)$$

From Figure 5(b), we can see that the noise voltage is actually of a random nature, and that the triangular approximation is for mathematical facility, but nevertheless its RMS value should be very close to the actual RMS value of the more random noise wave. The total area under the error wave of Figure 5(b) represents the total noise power. If now we assume this error wave to be completely random, then the noise would be evenly distributed throughout the frequency spectrum and the noise power in the pass band of a low pass filter would be directly proportional to the bandwidth of the filter, or

$$N_f^2 \propto f_{co} \qquad (2)$$

where N_f represents the noise voltage in the filter band and f_{co} is the upper cut off frequency of a low pass filter. Also, if the pulse rate of this error wave were to be increased, say doubled and the step height left the same, the area under the wave would not be changed, and the noise power would be the same, however the frequency spectrum of this noise would be twice as great, with the result that only one half as much noise power

would lie in the same band from 0 to f_{co}.

$$\therefore \quad N_f^2 \propto \frac{1}{f_p} \qquad (3)$$

Combining the effects of equations (1), (2), and (3), we get the effective noise voltage in the pass band of a low pass filter to be:

$$N_f = k \frac{\sigma f_{co}^{1/2}}{\sqrt{12}\, f_p^{1/2}} \qquad (4)$$

For maximum signal to noise ratio, we would want our signal to have an amplitude equal to $\frac{f_p \sigma}{2\pi f_s}$. Thus our ratio becomes:

$$\frac{S}{N_f} = \frac{\dfrac{f_p \sigma}{\sqrt{12}\, 2\pi f_s}}{k \dfrac{\sigma f_{co}^{1/2}}{\sqrt{12}\, f_p^{1/2}}} = K \frac{f_p^{3/2}}{f_s\, f_{co}^{1/2}} \qquad (5)$$

This figure is in exact agreement with that obtained by de Jager (3) by a slightly different approach.

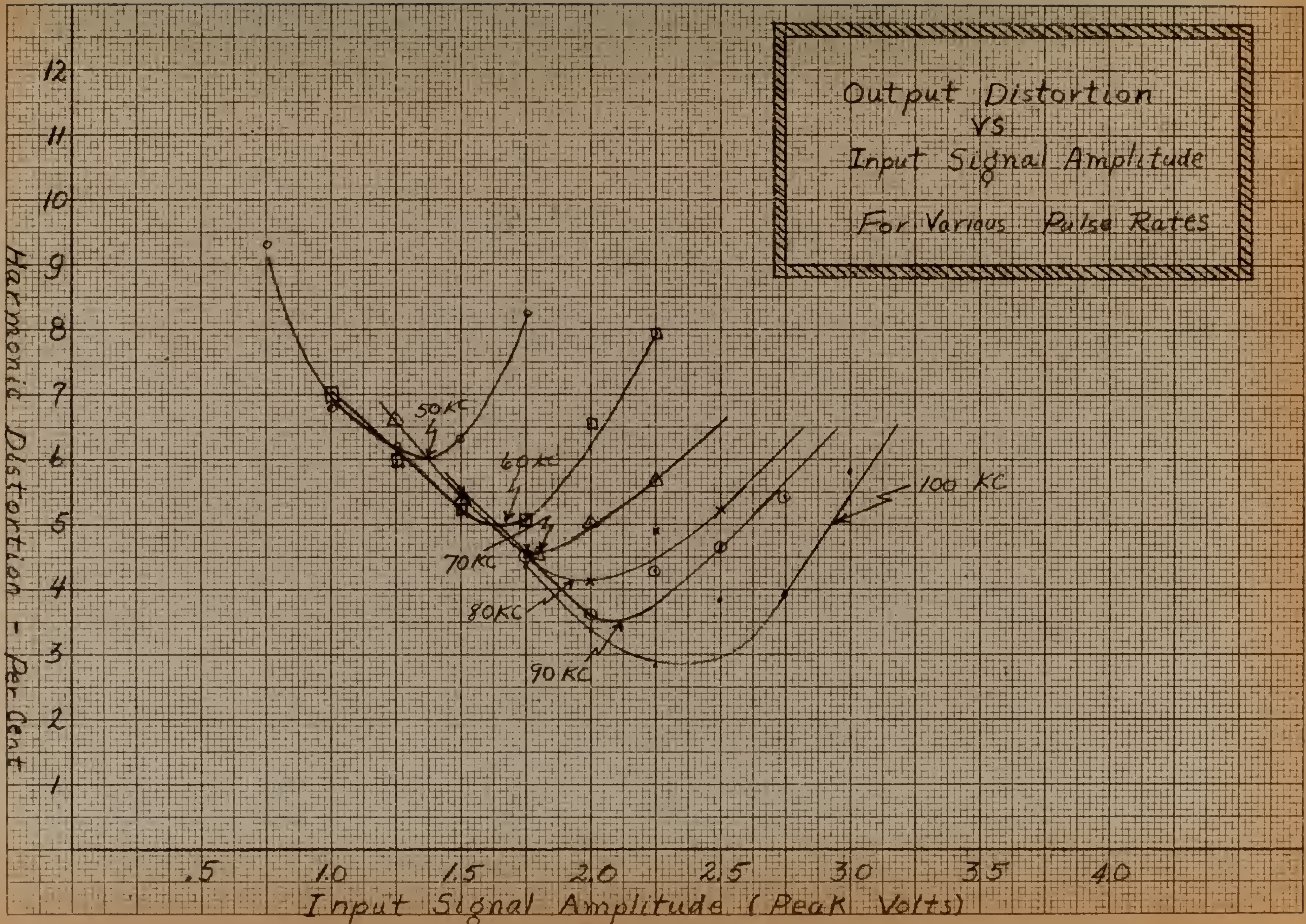
Harmonic Distortion in Delta Modulation

In the foregoing development the assumption of a purely triangular error voltage will not lend itself to any computation of purely harmonic distortion, because it is obvious that a chain of triangular waves of period T_p has no frequency components in the pass band of a low pass filter from 0 to f_{co}, where f_{co} is small compared to f_p. Rather, the harmonic distortion which results in this system is due to the harmonic irregularities of this triangular waveform. These harmonic irregularities will be of a completely different nature for each signal amplitude and any assumption as to their waveform or periodicity cannot be readily determined. To obtain a figure for purely harmonic distortion, it was decided to build up an experimental system and measure harmonic distortion of the image voltage as a function of pulse rate. A single integration, uni-polarity pulse

system, shown schematically in Figure 15 of the following chapter was constructed. This system is of original design and although it may not produce the most optimum performance, it is believed that its performance will be basicly characteristic of the delta modulation process in general. The input to the system was a sine wave of 1000 cps. This input signal amplitude was varied from a very low value upward in quarter volt steps with output harmonic distortion computed for each input signal amplitude to determine the optimum distortion figure. This was done for pulse rates of 100, 90, 80, 70, 60, and 50 kc. Step height and pulse width remained constant throughout. The stepped image was analyzed by use of a Hewlett Packard Wave Analyzer, Model 300A. The fundamental and first three harmonics were measured on this instrument. This simulated a low pass filter with perfectly flat response out to a cut off frequency of 4kc. The RMS value of the three harmonics was determined and this expressed as a per cent of the RMS of the fundamental.

Figure 10 is a plot of the principal results of these measurements. Only the points on either side of the region of optimum harmonic distortion are shown. Distortion increases for signal amplitudes greater than the amplitude which produces optimum distortion due to the onset of diagonal distortion, while distortion increases for amplitudes less than the optimum amplitude due to reduced signal amplitude with approximately the same noise power being present in the filter band.

Figure 11 is a plot of sampling rate versus optimum harmonic distortion and produces probably the most illuminating fact of the investigation. This plot shows an almost inverse linear relationship between harmonic distortion and sampling rate. The relationship previously given for signal


Output Distortion
VS
Input Signal Amplitude
For Various Pulse Rates
50KC
60KC
70KC
80KC
90KC
100 KC
Harmonic Distortion - Per Cent
12
11
10
9
8
7
6
5
4
3
2
1
.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
Input Signal Amplitude (Peak Volts)

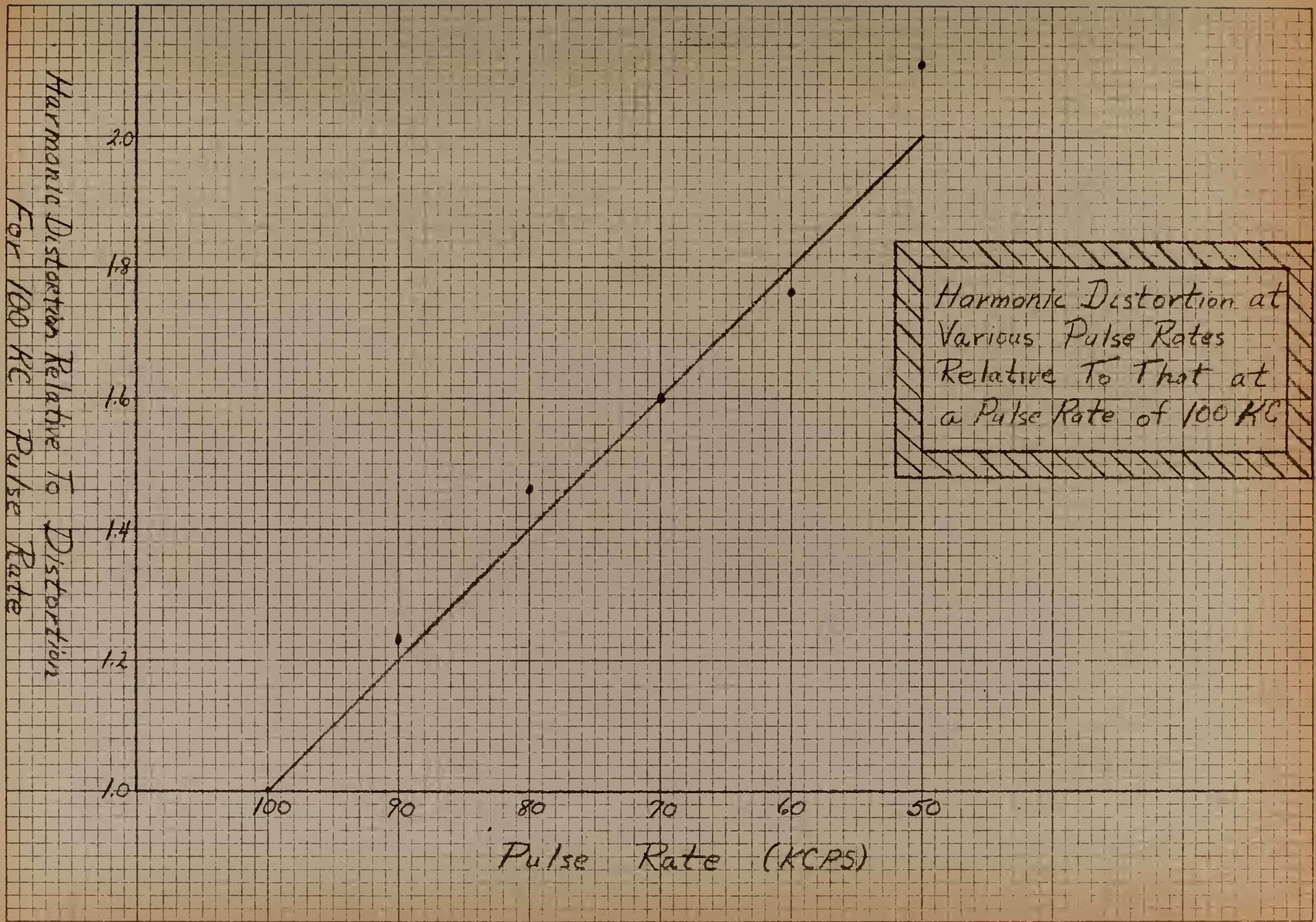

Harmonic Distortion at
Various Pulse Rates
Relative To That at
a Pulse Rate of 100 KC
Harmonic Distortion Relative To Distortion
For 100 KC Pulse Rate
2.0
1.8
1.6
1.4
1.2
1.0
100
90
80
70
60
50
Pulse Rate (KCPS)

Figure 11

to noise ratio might lead us to expect an inverse three halves power relationship for this distortion figure rather than the inverse first power relationship obtained. Two factors were observed to account for this discrepancy. First, the non-signal voltage components present in the complex wave analyzed were far from completely random as assumed in the idealized signal to noise ratio development. Rather, they existed in definite and distinct frequencies in the pass band, both on and in between the harmonic components measured. This non-random noise voltage observation agrees with the observation of de Jager (3) in some of his experimental work. Secondly, the amplitude of the non-signal voltage components decreased noticeably with increasing frequency, so that the low end of the spectrum appeared to have a larger percentage of the noise power than would otherwise be expected even under a complete random frequency distribution.

It is believed that this inverse first power relationship is more nearly the true-to-life figure than the theoretical inverse three halves power relationship which might be expected from the idealized computation of the signal to noise ratio figure.

In general we can therefore state that harmonic distortion for a single integration delta modulation system can be expressed as:

$$D = \frac{k}{f_p}$$

where D is harmonic distortion in per cent, and k is a constant characteristic of a particular system. For the system used in this work, k takes on a value of 3.3×10^5.

Figure 12 is a plot of pulse rate versus signal amplitude at which optimum harmonic distortion occurred. This plot agrees very well with

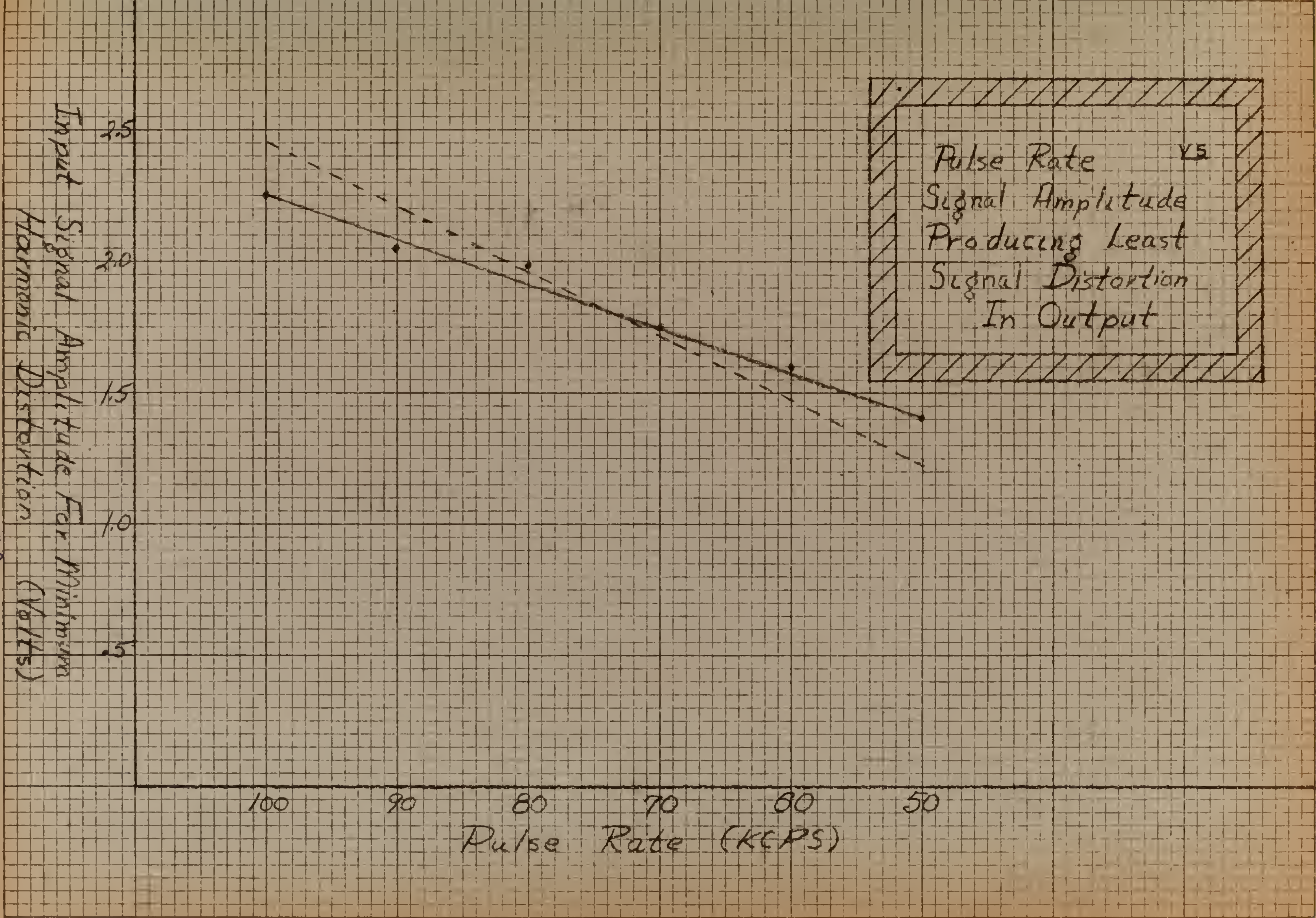

Pulse Rate vs
Signal Amplitude
Producing Least
Signal Distortion
In Output
Input Signal Amplitude For Minimum Harmonic Distortion (Volts)
2.5
2.0
1.5
1.0
.5
100
90
80
70
60
50
Pulse Rate (KCPS)

Figure 12

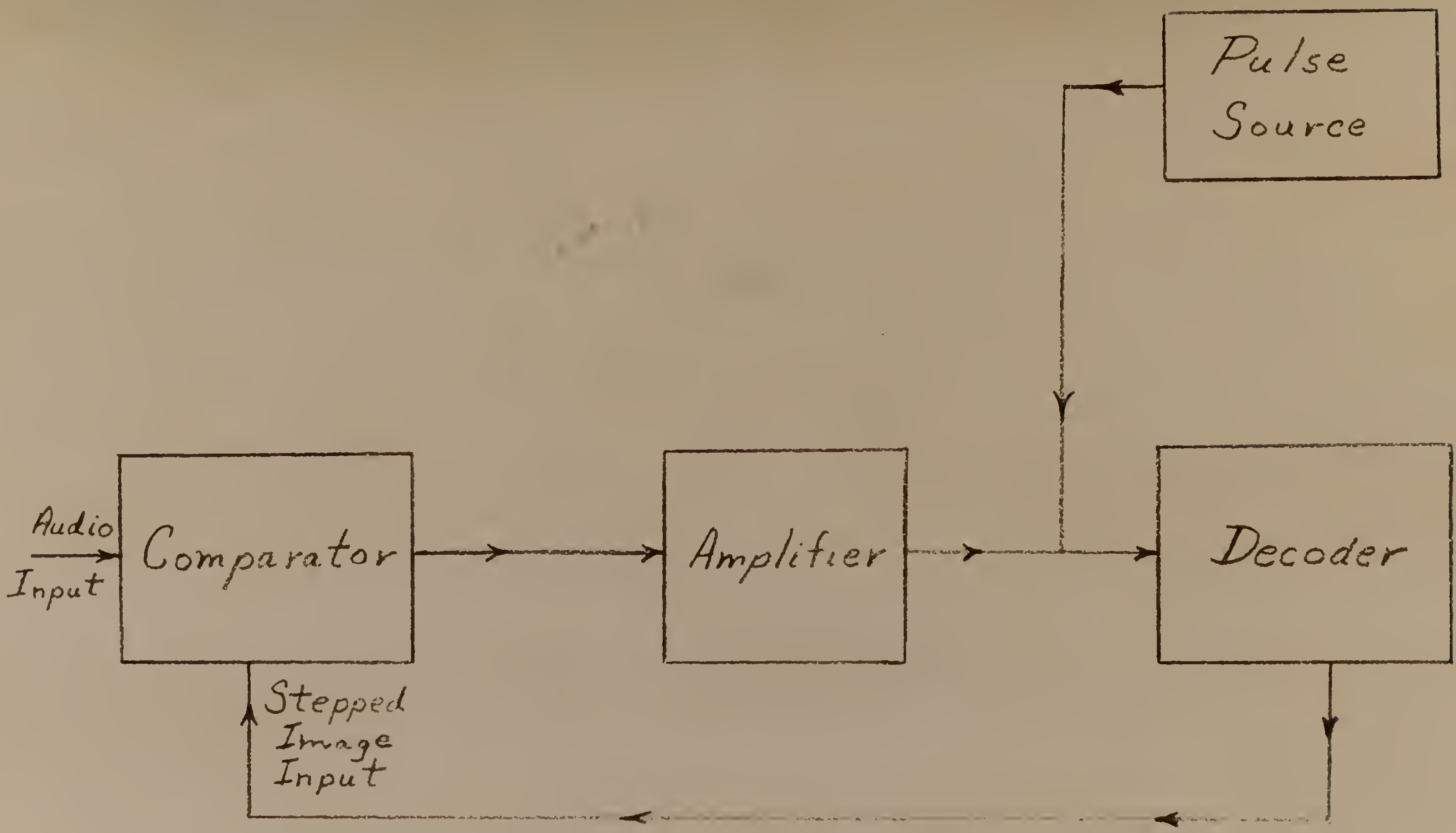
what we would expect from theory. The solid line represents actual observed values, while the dotted line represents the theoretical relationship.

CHAPTER VII

Two Experimental Systems And Concluding Remarks

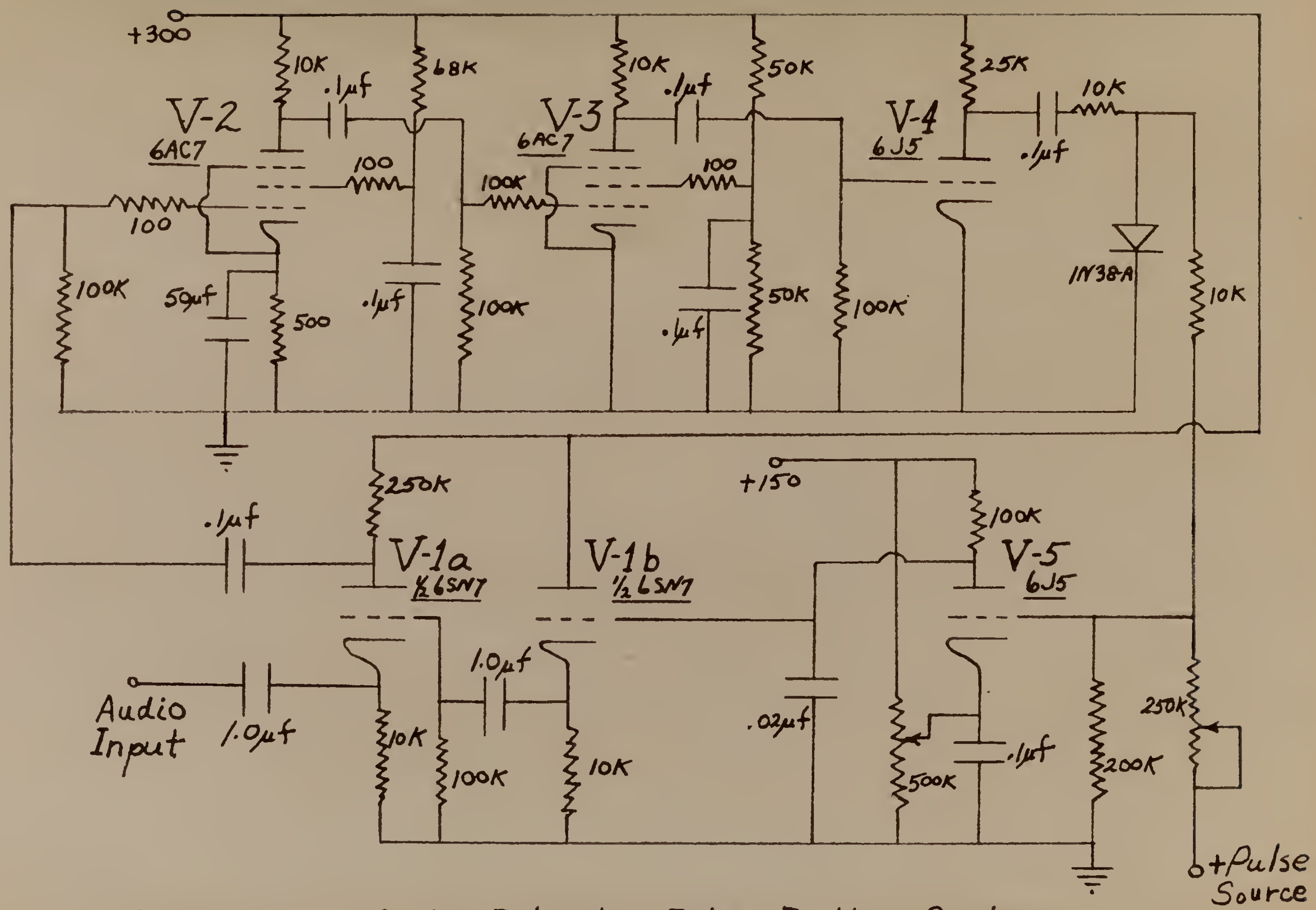
To add to the completness of this work, two experimental delta modulation systems of original design were constructed. A uni-polarity pulse system was constructed in the laboratories of the U. S. Naval Postgraduate School, and a dual polarity system was built up at the Federal Telecommunications Laboratories of Nutley, New Jersey, while the author was assigned to that concern during one term of graduate schooling.

Figure 13 is a block diagram of the uni-polarity pulse system. Figure 14 is the complete schematic of this same system. Two inputs are injected on the grid and cathode of the comparator tube V-1a. Since the grid and cathode have different gains, equal amplitude inputs at the two points do not cause a zero output from this tube. For the circuit constants used here it was found that this type comparator generated an image signal 1.08 times as large as the audio input signal which was applied to the cathode. The error voltage output of the comparator V-1a is amplified through the tubes V-2, V-3, and V-4. The output of V-4 will be squared up due to the limiting action of these three tubes and is used to gate the decoder tube V-5 on or off. The pulse source, a standard pulse generator, delivers positive pulses only to the grid of V-5 at a constant rate. Only when the output of V-4 is positive do these pulses produce conduction in the decoder tube V-5. The decoder is basicly a sawtooth generator, held cut off by the voltage divider action of the 500 kilohm potentiometer in the cathode circuit. With the output of V-4 negative, the positive pulses are not of sufficient amplitude to cause V-5 to conduct and the .02 μf



Uni-polarity Pulse Delta System

Figure 13

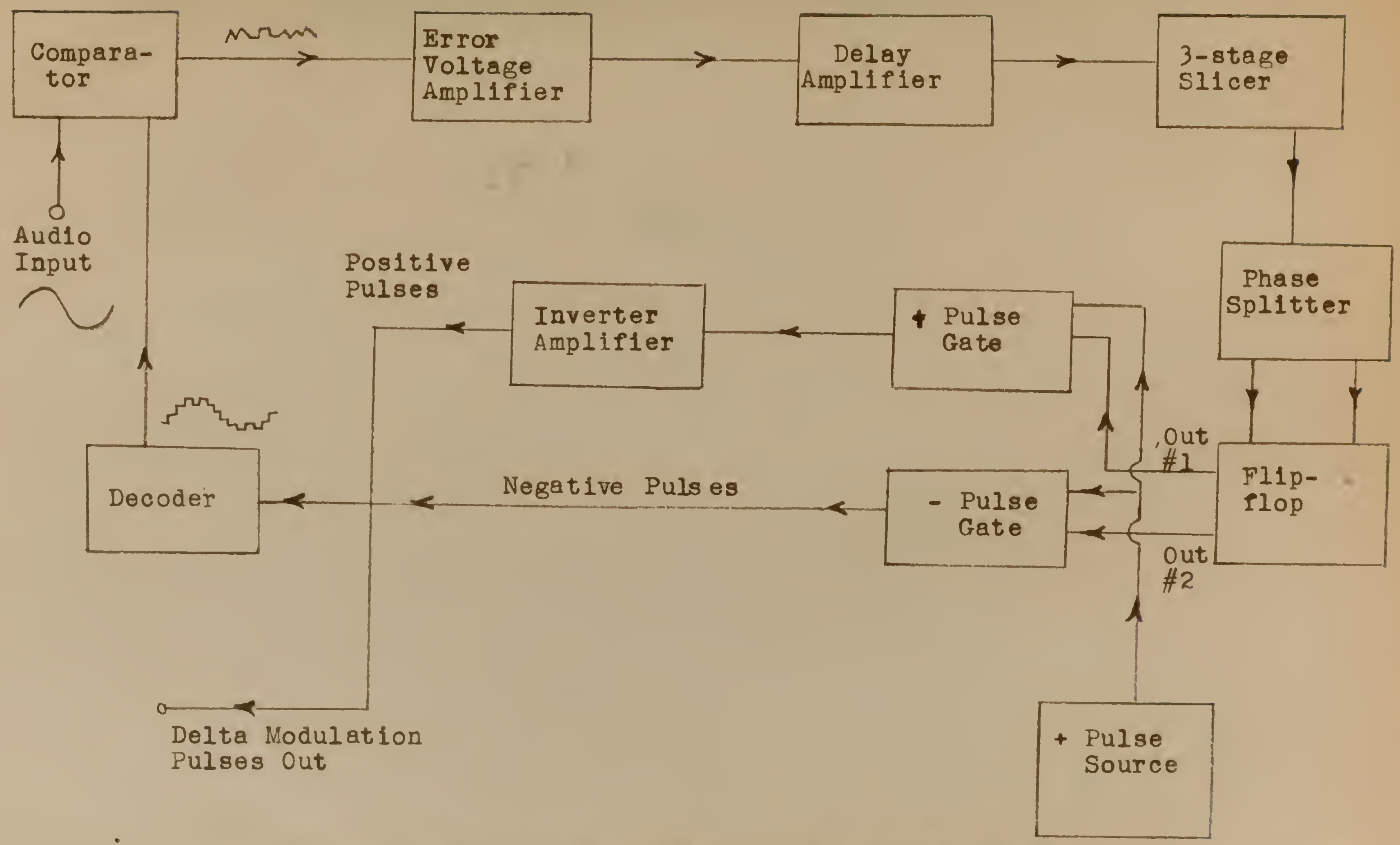


Schematic of Uni-Polarity Pulse Delta System

Figure 14

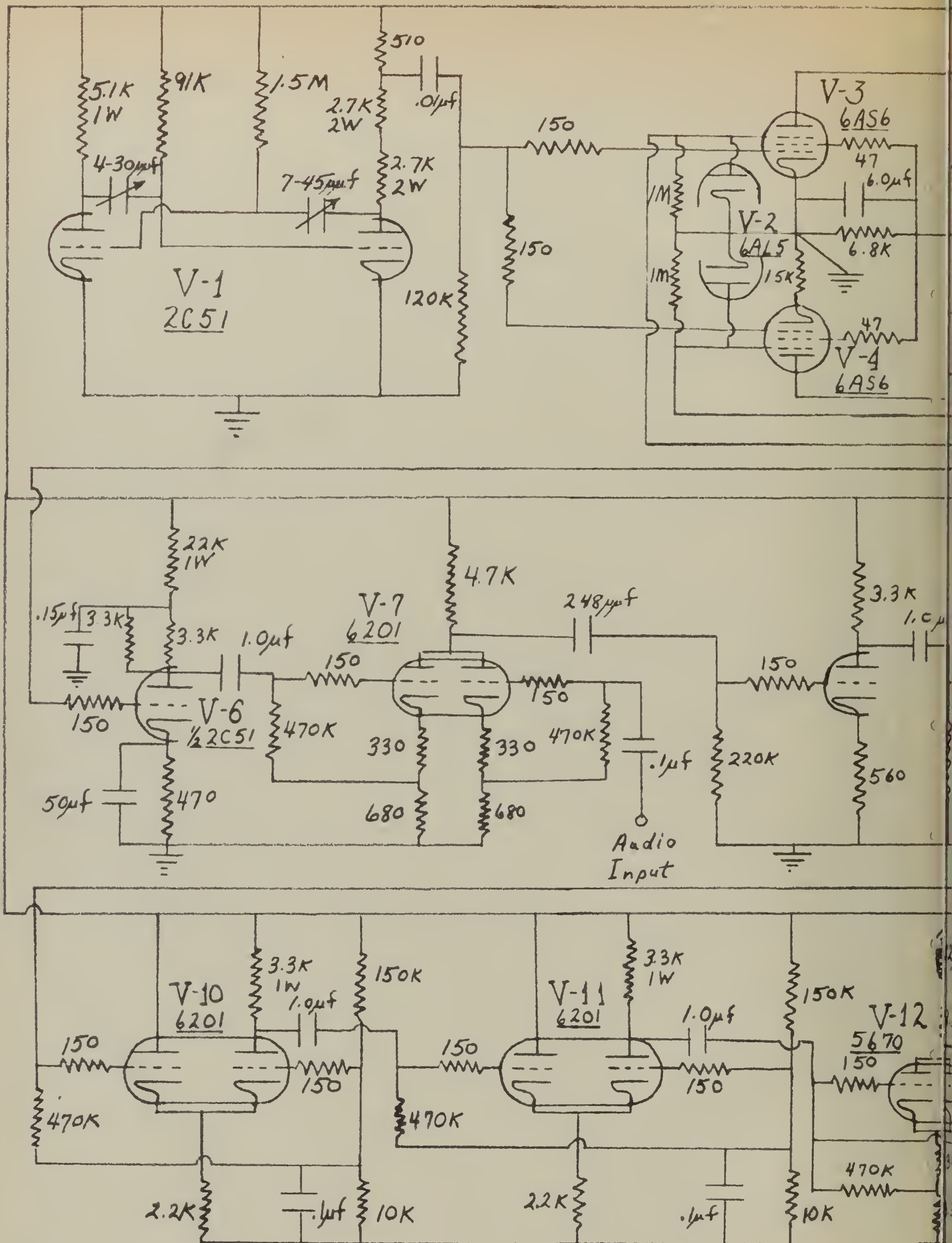
decoding capacitor across V-5 charges up toward 150 volts. When the output of V-4 is positive the positive pulses can set the decoder tube V-5 into conduction and this decoding capacitor discharges through the tube producing a step in the opposite direction. The plate of this decoder will operate about some mean DC level which should be adjusted to be midway between the plate supply voltage and the value of plate voltage resulting upon the application of a positive pulse. A DC input signal results in alternate pulses setting V-5 into conduction and the image signal is a sawtooth waveform with a frequency equal to one half the pulse frequency. The output of this decoder is fed through the cathode follower V-1b to the grid of V-1a for comparison with the audio signal on the cathode.

The second experimental system built was a dual polarity system, the block diagram and schematic of which are shown in Figures 15 and 16 respectively. In this system, the error voltage out of the comparator V-7 is utilized after amplification to control the state of a flip-flop, tubes V-14 and V-15. The two plate outputs of this flip-flop are sent to the suppressor grids of the two gate tubes V-3 and V-4 and thus control the conduction state of the gates. Each gate receives continuous positive pulses on their control grids from the pulse source V-1. As only one of the two gate tubes can be in a conduction state at any one time, the grid of the decoder tube V-5b will receive either a positive or a negative pulse depending upon which gate tube is in a conduction state to pass the grid pulse. Although each gate tube delivers negative pulses as its output, the output of gate tube V-3 is sent through inverting amplifier V-5a to produce the positive pulses. These positive and negative pulses on



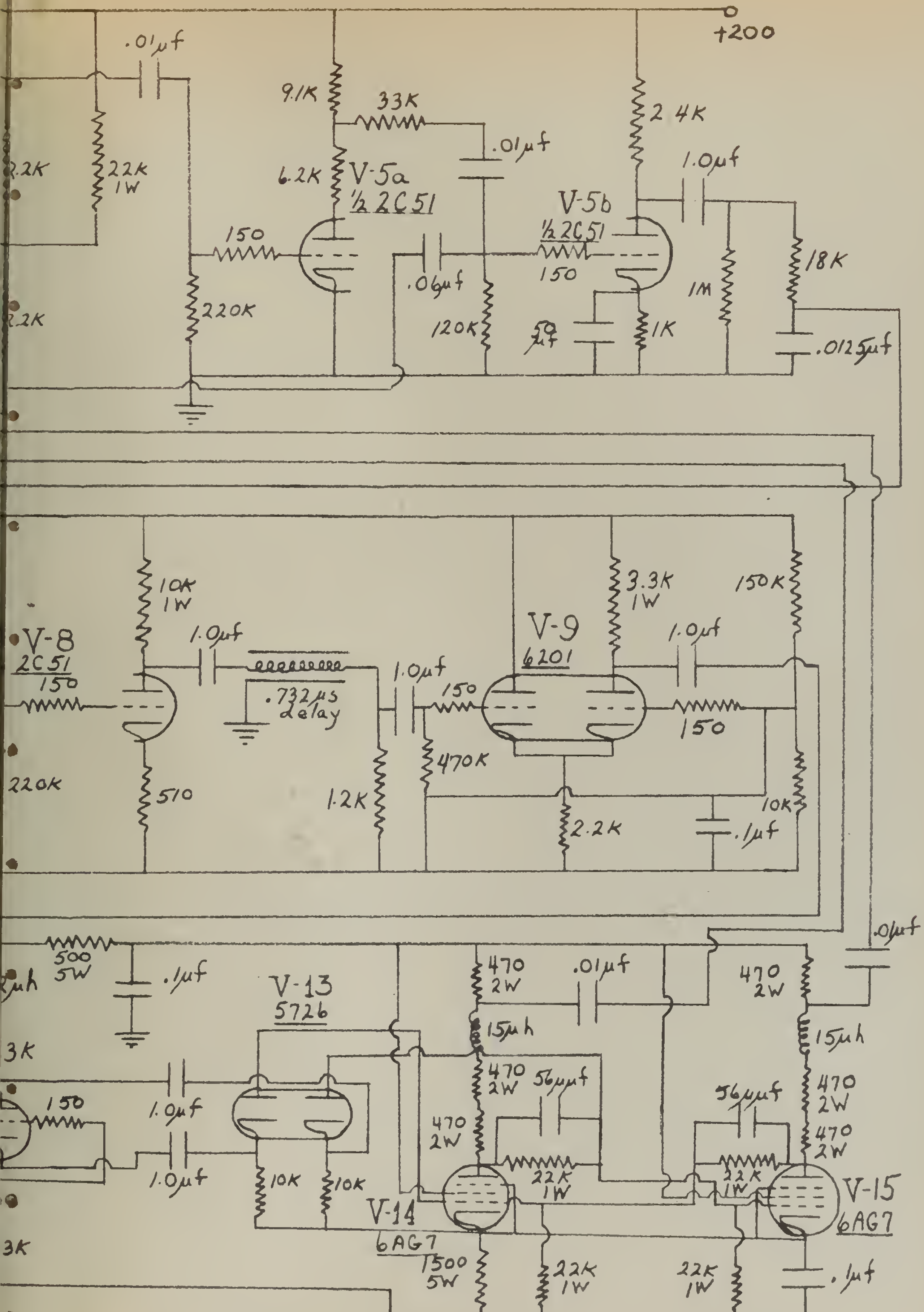
Block Diagram of Dual Polarity Pulse Delta System

Figure (15)



Schematic of a Dual Polarity Puls

Figure 16

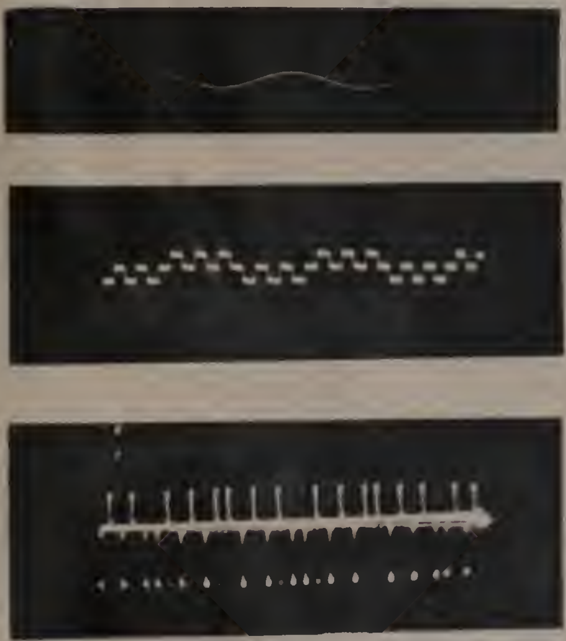


Delta System

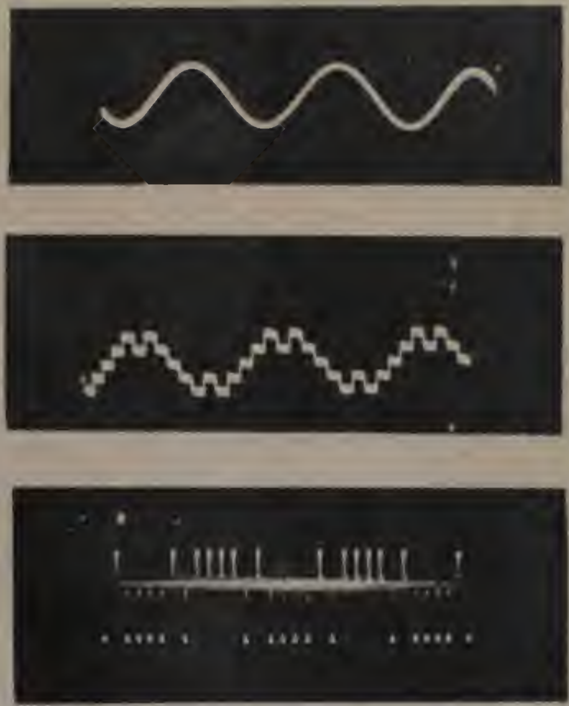
the grid of decoding tube V-5b produce a stepped output across the .0125 integrating capacitor in the plate circuit of V-5b. The 18 kilohm resistor and this capacitor constitute the integrating circuit which is the heart of the decoder. The integrating action of these components on the plate pulses generated as a result of the delivered grid pulses result in the stepped image desired. This stepped output is passed through the decoupled stage V-6, inserted to counteract the principal time constant of the decoder which is relatively small compared to the audio frequencies involved. The comparator tube V-7 takes this stepped image plus an audio input and produces the error voltage which determines the state of the flip-flop which in turn determines the polarity of the next applied pulse. The delay amplifier V-8 is inserted to prevent the flip-flop from being triggered until after the termination of any pulse producing an initiating change. Slicers V-9, V-10, and V-11 provide the necessary amplification of the error signal and also square it up so as to provide adequate triggering for the flip-flop. The flip-flop is triggered by the application of negative impulses to the grids. To provide this requisite triggering action the phase splitter V-12 was inserted to provide a negative going impulse to one of the triggering diodes of V-13 regardless of the polarity of the error signal at the output of V-11.

Figure 17 is a group of waveform photographs of the dual polarity system in operation. Photograph (a) is the decoder output to the comparator for a DC signal input, plus the grid pulses taken from the grid of tube V-5b. It will be noticed that the alternate positive and negative pulses on the grid cause alternate positive and negative steps in the stepped output of the decoder. Photograph (b) shows a sine wave signal

Decoder Output and Delta
Modulation Grid Pulses
for DC Input.
(a)



Small Amplitude Sine Wave
Signal with Resultant
Decoder Output and Delta
Modulation Grid Pulses.
(b)



Larger Amplitude Sine Wave
Signal with Resultant Decoder Output and Delta
Modulation Grid Pulses.
(c)

Sine Wave Signal Producing
Saturation and Decoder Output.
(d)

Figure (17)

of small amplitude which was the audio input to the comparator. Below this is the resultant delta modulation stepped image from the decoder and below that the delta modulation pulses on the decoder grid which produced the stepped waveform. Photograph (c) shows the same sine wave signal increased in amplitude, the delta modulation stepped image from the decoder plus the delta modulation pulses from the grid of the decoder. Photograph (d) shows a sine wave signal of still larger amplitude which succeeded in completely saturating the system. Below this is the stepped image from the decoder. The pulses producing this are not shown but would consist of alternate groups of several positive and several negative pulses.

Concluding Remarks

Delta modulation can be seen to be a very worthwhile addition to the pulse modulation field. Combining the beneficial attributes of pulse code modulation with a more simplified circuitry it may well prove to be the more practical system. No attempt has been made to make an exacting comparison between the two systems; however, those interested in some of the comparisons now on record are referred to the articles of deJager (3), Libois (5) and Rose (7). A comparison of Figure 14 and 16 will clearly show how much the uni-polarity pulse technique simplifies the required circuitry.

The future of delta modulation is very promising. Not only could it be used on a large multi-channel scale, but due to its simplified circuitry might easily be adapted for use on a single channel basis over circuits or transmission paths too noisy for normal communication methods and too little used to warrant a more expensive installation. In this

line it might well find beneficial military as well as commercial applications. A one channel pulse delta system of the uni-polarity type, battery powered, could undoubtedly be built on a scale comparable to the familiar "walkie-talkie", providing in the field the transmission reliability and quality of pulse code modulation itself plus a certain amount of communications security. This would certainly be an application worthy of further investigation and consideration.

BIBLIOGRAPHY

1. Andrews, B. H. Methods of Modulation. Thesis, unpublished, U. S. Naval Postgraduate School, 1948.

2. Bennett, W. R. Spectra of Quantized Signals. The Bell System Technical Journal. Vol. 27, 446-471, 1948.

3. de Jager, F. Delta Modulation, A Method of P.C.M. Transmission using the 1-unit Code. Philips Research Reports. Vol. 7, 442-466, 1952.

4. Laboratoirie Central de Telecommunications, Paris, France. Note on Delta Modulation. Electrical Communications. Vol. 30, 70-74, 1953.

5. Libois, L. J. Un Nouveau Procede de Modulation Codee: la Modulation en Delta. L'Onde Elect. Vol. 32, 26-31, January, 1952.

6. Oxford, A. J. Pulse Code Modulation Systems. Proceedings of the Institute of Radio Engineers. Vol. 41, 859-865, July 1953.

7. Rose, A. E. Multiplex Transmission Systems, Time Division and Frequency Division. Thesis, unpublished, U. S. Naval Postgraduate School, 1953.

8. Sanger, F. M. Jr. Pulse Modulation Developments Leading to PCM. Unpublished report from Massachusetts Institute of Technology Electrical Engineering Seminar, 1948.

9. Schouten, J. F., de Jager, F., and Greefkes, J.A. Delta Modulation, A New Modulation System for Telecommunication. Philips Technical Review, Vol. 13, 237-268, March 1952.

10. Sears, R. W. Electron Beam Deflection Tube for Pulse Code Modulation. The Bell System Technical Journal. Vol. 27, 44-57, January, 1948.

11. Smith, B. D. Coding by Feedback Methods. Proceedings of the Institute of Radio Engineers. Vol. 41, 1053-1058, August, 1953.

12. Vail, C. R. Circuits in Electrical Engineering, Prentice-Hall, Inc., New York 1950. 441-473.

25279

Thesis
A716

Armstrong
A study of pulse delta modulation.